基于区块链与函数加密的隐私数据安全共享模型研究

李懿[1,2,3]，王劲松[1,2,3]，张洪玮[1,2,3]
1. 天津理工大学计算机科学与工程学院，天津 300384；
2. 智能计算机及软件新技术天津市重点实验室，天津 300384；
3. 计算机病毒防治技术国家工程实验室，天津 300457

摘要
区块链技术给数据共享中的数据确权、数据溯源、数据可信、数据可用等方面提供了新思路，但数据共享中的隐私数据安全仍面临许多挑战。首先回顾当前基于区块链的数据共享研究现状；然后提出一种隐私数据安全共享模型，通过函数加密技术对隐私数据进行加密，结合零知识证明技术生成相关计算的可信证明，实现“数据可用不可见”的安全可靠的数据共享。实验结果显示，该模型的共享时延及经济开销均在可接受范围内，证明了该模型的安全性和可行性。

关键词
区块链；函数加密；零知识证明；数据共享；隐私保护

中图分类号：TP309.2 文献标志码：A doi: 10.11959/j.issn.2096-0271.2022072

Research on privacy data security sharing scheme based on blockchain and function encryption

LI Yi[1, 2, 3], WANG Jinsong[1, 2, 3], ZHANG Hongwei[1, 2, 3]
1. School of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China
2. Tianjin Key Laboratory of Intelligence Computing and Novel Software Technology, Tianjin 300384, China
3. National Engineering Laboratory for Computer Virus Prevention and Control Technology, Tianjin 300457, China

Abstract
Blockchain technology has provided new ideas for data validation, data traceability, data trustworthiness, and data availability in data sharing, but privacy data security in data sharing still faces many challenges. Firstly, the current status of blockchain-based data sharing research was reviewed. Then a secure sharing model of privacy data was proposed. By encrypting the privacy data through function cryptography, and generating the proof of computational correctness through zero-knowledge proof technology, a secure and reliable data sharing with “data available but not visible” was realized. The experimental results show that the sharing delay and economic overhead of the model are within the acceptable range, which demonstrates the security and feasibility of the model.

Key words
blockchain, functional encryption, zero knowledge proof, data sharing, privacy protection

0 引言

现代生活中，大部分领域实现了数字化，如商品、商业、服务等。通信技术、硬件技术的发展加速了大数据时代的到来，数据成为新的生产资料和价值高地。大量的个人数据被移动设备采集并存入云端。依赖于大量数据的个性化服务（如商品推荐、健康监测）随之出现。与此同时，公众对隐私数据的关注度日益增长，加上相继出台的隐私数据保护法规，各大公司和服务提供商正积极探寻既能维持用户信任又能合法合理收集数据的方法。虽然用户可以使用简单的端到端加密算法来直接提升数据安全性，但数据加密也将导致现在被广泛接受的便利服务消失，这是因为服务提供商无法获得并分析原始数据。因此，隐私计算的概念应运而生，其主要目标是实现“数据可用不可见，数据不动模型动”，让用户将数据保存本地或者加密后上传至云端，服务提供商只获取数据处理结果，从而保护原始数据。常见的隐私计算方法有安全多方计算、零知识证明、可信执行环境等。此外，还有一些是能够在保护隐私的同时不降低数据可用性的加密算法，函数加密就是其中的一种。与同态加密技术类似，函数加密允许数据使用者直接从密文中获取信息，并且数据拥有者可以精确控制数据使用者从密文中获取的信息量，这给数据安全共享带来了新的可能性。因此，本文提出了基于区块链与函数加密的隐私数据安全共享模型，旨在实现面向隐私保护的数据安全共享平台。

1 研究现状

传统中心化数据共享具有单点故障、数据易丢失、易被篡改、隐私数据难保护等问题。近年来随着区块链技术的兴起，研究热点已经转移到去中心化数据共享。Chen J C等人[1]利用区块链难以篡改的特性，将系统中的每次数据共享记录及相关信息记录在区块链中，并利用智能合约充当数据交换的媒介。但是由于区块链具有数据透明的特点，直接将原始数据暴露在链上会导致数据隐私难以通过该方案共享。Liang X P等人[2]利用联盟链创建了一个以用户为中心的健康数据共享模型，用户的可穿戴设备和医疗设备收集到的数据会同步存储到云端，并且为了保证数据的完整性，云端会把数据的检索记录和验证结果传至区块链进行保存，还利用基于树的方法提高平台的可扩展性和性能。但是其假设云是可信的，这大大减弱了该模型的安全性。Theodouli A等人[3]提出基于区块链的电子健康信息共享平台，其假设云服务器是恶意的，并通过对比链上数据的哈希值，确保恶意云服务器没有篡改数据，还给用户提供了完善的页面，帮助用户快速获知自己数据的访问情况，起到审计和监督的作用。但是其并未考虑恶意服务器泄露数据的风险，并且其假设网络由可信实体维护，这使得该方案难以落地。Gordon W J等人[4]分析了基于区块链构建医疗数据共享架构时需考虑的安全风险，但是其并未给出具体的解决措施和架构。Azaria A等人[5]提出去中心化的电子病历管理系统，用户可以通过智能合约自定义访问控制策略，并与不同的机构共享自己的电子病历。此外，作者还设计了激励系统来鼓励权威机构积极参与，维护网络稳定。Yang H等人[6]借助区块链和属性基加密（attribute-based encryption，ABE）实现安全电子健康数据共享，用户将能够访问数据的属性条件记录在链上，只有拥有相关属性的人才能

向云服务器发起数据访问请求。Cao S等人[7]提出了云链结合的安全电子医疗数据共享系统，其核心思想是将每个数据操作写为公有链中的一笔交易，并在数据上传时将数据的哈希值与签名存在链上，从而保证数据的正确性和完整性。刘彦松等人[8]利用属性加密和同态加密构建了保护隐私的链上数据交易平台，利用同态加密对密文进行计算，从而避免泄露原始数据，但是其并未给出该系统的性能指标，并且同态加密的计算开销大，复杂计算难以在链上实现。Yu K P等人[9]利用智能合约和属性基加密，实现了工业互联网的数据共享平台，域管理员负责制订域安全和访问策略。其中拥有与访问策略相匹配的属性的用户，可以从边缘/云服务器获取中间解密参数，并且支持用户属性的更新和撤销。Zhang Y等人[10]和Wörner D等人[11]都通过在比特币中写入自建协议，配合链下数据解析脚本，构建IoT设备数据共享平台。这种方法利用比特币的稳定性来保证共享服务的稳定性，并利用假名技术保护用户隐私。但是比特币的性能不高，难以满足IoT的低时延需求。Özyilmaz K R等人[12]利用智能合约构建了IoT传感器与数据使用者的数据共享平台，但是IoT设备通常不具有运行区块链的能力，并且该研究同样没有考虑隐私数据。Shafagh H等人[13]利用区块链实现了去中心化的访问控制和数据管理，并通过传统访问控制方法来保护隐私数据，但是其不支持访问策略的更新。Pan J L等人[14]利用智能合约控制IoT设备从边缘服务器获取数据，边缘服务器通过读取链上数据来判断IoT设备访问是否合法。张召等人[15]对现有基于区块链的数据共享模型进行了分析，指出了该架构中存在的问题，并提出了链上数据完整性存证、链下实际数据传输的数据共享模型，但是并未解决隐私数据问题，且未给出实验证明。

通过对比现有数据共享研究不难发现，尽管在该领域已有一些研究成果，但是还存在诸多可以改进的地方，例如未考虑隐私数据、将云服务器假设为完全可信等。有些问题难以单独使用区块链或云来解决。为此，本文提出了基于区块链与函数加密的隐私数据安全共享模型，实现了更加高效、安全的数据共享。

2 相关技术

2.1 区块链及智能合约技术

区块链是一个公开、分布式、难以篡改的账簿，其最早在中本聪于2008年发布的比特币白皮书中被提出，其被作为支撑比特币系统的底层技术。区块链通常由运行相同或相近版本的客户端软件的节点通过对等网络组成，网络中传递的能够引起系统状态变化的消息被称为交易。整个对等网络中不存在中心化节点，所有节点间的地位及权利都是相等的，只有运行提前商定的共识算法，网络中的所有节点才能够对整个网络的状态变更达成统一。常见的共识算法包括工作量证明（proof of work，PoW）、权益证明（proof of stake，PoS）、权威证明（proof of authority，PoA）等[16]。总体来说，区块链具有多种性质，这里仅列出与本文关系密切的特性。

● 透明性。区块链中的任何节点都可以随时获取区块链的所有状态，包括历史状态与当前状态，并且所有的历史交易与当前正在网络中传播的交易都是透明的。

● 永久存储。区块链会永久记录所有造成状态变更的交易，一旦区块被添加并确认，其内容就是不可更改的，这也是区块链技术非常突出的特性。

● 去中心化。区块链网络中不存在中心节点，这意味着只要交易自身是正确的，其一定会被忠实地执行。并且由于不存在中心节点，没有任何单独实体能够控制整个网络，除非所有节点同时宕机，否则区块链网络将永远存在。

智能合约是区块链发展过程中极其重要的里程碑，简而言之，智能合约可以被看作预先定义好的一段程序代码，代码通过交易的形式被存储在区块链中，并得到网络中各参与节点的共识。智能合约通常由Solidity语言编写，只能被区块中的交易触发执行。智能合约是否被正确执行，由所有区块链节点共同验证。结合区块链自身的加密货币属性与共识协议中的执行激励，智能合约可以实现复杂的交易和业务逻辑。以太坊是第一个支持图灵完备智能合约的区块链，其设想最早由Vitalik Buterin于2014年提出。以太坊存在两种账户：外部账户和合约账户。外部账户由唯一对应的公钥和私钥控制，能够通过私钥进行签名并发起交易，其地址由公钥计算得到。合约账户与外部账户最大的不同在于没有私钥能够控制合约账户，并且合约账户能够存储代码，其账户地址由固定算法直接算出。与比特币一样，每个交易都由输入方和接收方构成，外部账户可以通过使用私钥签名的交易进行转账、合约创建、合约调用等一系列操作。智能合约的行为完全取决于预设代码及外部账户调用传入的参数。

2.2 函数加密技术

函数加密是在Shamir A[17]提出的身份基加密（identity-based encryption，IBE）和Goyal V等人[18]提出的属性基加密的基础上发展而来的一种新型公钥加密算法，最早由Boneh B等人[19]于2011年正式提出。它从两个方面拓展了传统公钥加密算法体系：一是其支持访问控制，只有满足特定条件的人才能解密；二是其允许通过对密文进行选择性的计算直接得到计算结果。

具体来说，传统公钥解密中往往只有解密成功或者解密失败两种结果，也就是说，解密者要么获取所有明文信息，要么不能获得任何信息。但函数加密允许加密者精确控制透露给解密者的信息量。例如，使用计算平均值的函数密钥对加密数据进行解密时，只能获取该加密数据的平均值，而不会获得额外信息。函数加密中每个函数F的密钥都与密钥空间K中的某个k值相关联，该密钥由可信权威机构利用全局主密钥生成。当消息空间X中的数据x被可信权威机构生成的主公钥加密后，使用与k相关联的密钥sk_k解密后的结果为$F(k,x)$。一个函数加密方案包含5个算法：Setup输入安全参数生成后续需要的素数群；Master Key Generation创建主公钥和主私钥；Function Key Derivation以主私钥和具体函数F为输入，生成sk_k用于获取函数结果；Encryption用于加密数据，生成密文；Decryption以sk_k和密文为输入，获取$F(x)$的解密结果。

2.3 零知识证明技术

零知识证明技术是一种加密技术，用于证明某个断言，而不具体透露其他任何信息。虽然其出现早于区块链，但却因为区块链被广为关注。由于区块链的透明性让网络中的用户可以随意访问所有

数据，容易造成隐私数据泄露等问题。零知识证明的特性使其成为实现区块链隐私保护的重要方法。零知识证明包含以下3个特性。

● 完整性：诚实的证明方产生的合法证明一定会通过验证。

● 可靠性：证据在计算上是不可伪造的。

● 零知识性：验证者除证据本身外不能获取任何额外知识。

零知识证明主要分为两类：交互式零知识证明和非交互式零知识证明。交互式零知识证明是指验证方多次向证明方的声明提出挑战，双方需要多次通信，直至验证方认为正确概率大于某个阈值时表示认可。同时，值得注意的是，在交互式零知识证明中，证明方只能同时向一方证明。非交互式零知识证明是指证明方单方计算证据，随后任何验证方都可以随时快速验证证据的正确性，因此现在广泛使用的都是非交互式零知识证明。非交互式零知识证明的主要参与方包含3个：Generator以安全参数k为输入，输出另一个安全参数pp；Prover生成证据π，用于证明某个声明的真实性；Verifier能够快速验证证据的真实性。

3 系统架构

3.1 设计目标

本文模型旨在达到如下设计目标。

● 数据拥有者可控的数据共享。针对数据易复制、难管控、难确权等问题，本文模型旨在让数据拥有者对数据拥有完全的控制权。本文模型具有原始数据不外露、云存储数据防篡改、访问控制、安全共享等特点。

● 隐私保护。为了在保护原始数据的前提下实现数据共享，本文模型旨在结合密码学和区块链技术，实现“数据可用不可见”、结果可验证，从而保证隐私数据的安全性。

● 可靠性验证。使用零知识证明技术，保证数据处理结果的可靠性，消除由原始数据不可见带来的数据处理不可信风险，保护数据使用者的权益。

3.2 整体架构

本文结合区块链与云服务器，实现数据链上存证、链下存储的混合存储架构，并使用函数加密、零知识证明等技术实现隐私保护以及可验证结果的安全数据共享，具体架构如**图1**所示。本文模型包含以下主要角色：可信机构（trusted authority，TA）、区块链（blockchain，BC）、云服务提供商（cloud service provider，CSP）、数据拥有者（data owner，DO）、数据使用者（data user，DU）。

各个角色具体介绍如下。

● 可信机构：负责全局初始化设置，本文模型中使用的函数加密及零知识证明的初始化参数设置都由TA生成，并且TA负责函数加密中私钥的生成。本文假设TA完全可信。

● 区块链：负责存储加密数据摘要、验证零知识证明的正确性、发布DO的定价及函数、接收DU的数据请求以及最终的自动支付流程，上述功能通过智能合约实现，合约中包含相关事件，用于传递消息。

● 云服务提供商：负责存储数据拥有者上传的密文，并向购买后的数据使用者提供密文。

● 数据拥有者：负责存储数据、加密数

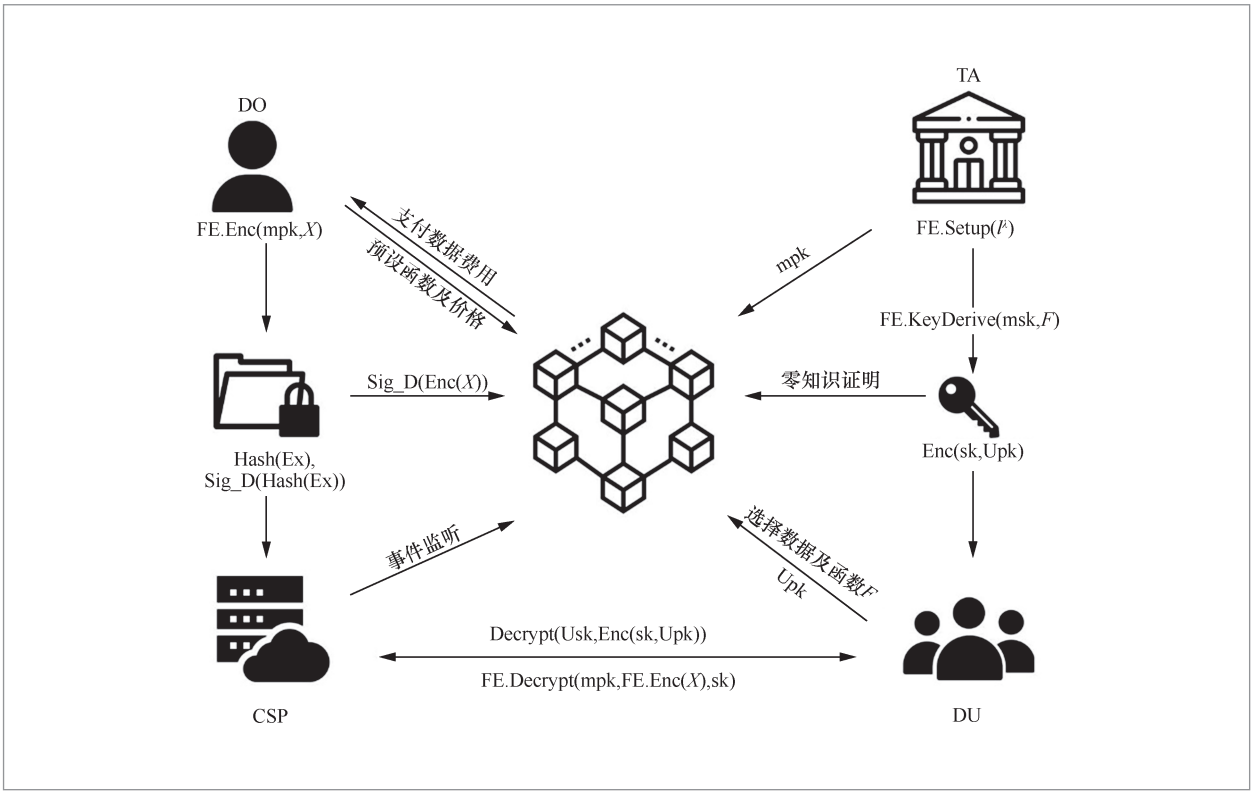

图1 整体架构及各角色间交互

据、指定使用函数并定价。

● 数据使用者：金融机构、科技公司、政府机构等具有私人数据使用需求的用户。

本文模型共涉及3种密码学算法，分别为函数加密、零知识证明和非对称加密。其中函数加密的相关符号均以FE开头，将在第4.1节详细论述。零知识证明的生成方法将在第4.2节详细论述。这里介绍非对称加密部分，即**图1**中的Hash(Ex)、Enc(sk,Upk)、Sig_D(Hash(Ex))、Decrypt(Usk,Enc(sk,Upk))。Hash(Ex)表示对函数加密后的密文Ex进行哈希计算。Sig_D(Hash(Ex))表示对密文哈希值进行数字签名，用于防止数据被篡改。Enc(sk,Upk)表示使用用户公钥Upk对派生出的密钥sk进行加密。Decrypt(Usk,Enc(sk,Upk))表示使用用户私钥Usk对加密后的sk进行解密，获取原始sk。

3.3 运行流程

本文模型的详细运行过程如下。

步骤1： 全局初始化。首先由TA设置全局安全参数，生成系统主公钥和主私钥，并将系统主公钥公开在区块链中。

步骤2： 数据加密及上传。DO从链上获取主公钥，并运行函数加密算法后，得到密文Ct。为了避免数据过大造成区块链节点负担过重，本文模型采用链上链下混合存储的方法。DO对密文Ct进行哈希摘要和数字签名，然后连同密文Ct一起上传到CSP，并将哈希摘要及签名通过交易发送至区块链进行存储，从而提高模型的性能及可扩展性，降低链上空间开销。同时，因为本文模型中假设CSP是诚实的，即CSP会诚实地执行用户的指令，

但是可能会对用户的数据产生好奇和恶意，这种存储方案也能避免CSP篡改数据。CSP会对链上事件进行监听，保证只有合法的DU能够拿到加密数据。然后，DO需指定能够作用在加密数据上的函数及对应的价格。由于本文模型采用了内积函数加密（inner product functional encryption，IP-FE），除DO预设函数外，DU可以自行上传加密数据处理函数，函数种类包括加权和、平均值，甚至简单的机器学习模型等。

步骤3：数据访问。当DU想要获取数据时，其并不直接获取原始数据，而是获取加密数据的函数处理结果，即$F(x)$。将传统数据共享中的共享原始数据改为共享数据的计算结果，实现数据的隐私保护共享。DU可以选择链上的预设函数或自行提供函数F，同时提供个人公钥Upk，该公钥用于对最终结果进行加密，最后通过交易调用智能合约。

步骤4：密钥生成。TA监听到链上事件后，从事件中获取DU购买的函数F及个人公钥Upk，运行密钥生成算法，对F及加密数据进行处理并生成sk。并通过零知识证明算法生成sk的有效性证明Proof，将Proof上传至链上进行验证。智能合约验证Proof的有效性后，会将步骤3中DU预存的资金发送给DO。如果验证未通过或在规定时间内TA未生成Proof，交易将会被取消并对TA进行惩罚。最后使用Upk加密的sk会被记录在合约中，相关事件被触发以通知DU获取sk。

步骤5：解密。DU首先利用Usk对链上获取的被Upk加密的sk进行解密。然后从CSP获取密文，并与链上的哈希值进行对比，验证文件是否被篡改，最后运行解密算法获取$F(x)$。

4 算法构造

下面详细阐述本文模型中使用的函数加密和零知识证明的具体算法设计及构造过程。

4.1 函数加密

本函数加密模型基于DDH（decisional Diffie-Hellman）假设构建。下面对DDH假设进行简单介绍。

$$(G,p,g)\leftarrow \mathrm{GroupGen}(l^{\lambda}) \tag{1}$$

其中，GroupGen为任意概率多项式时间算法，输入安全参数l^{λ}，产生一个三元组，其中G为p阶素数群，其生成元为g。那么(g,g^{a},g^{b},g^{ab})与(g,g^{a},g^{b},g^{c})是不可分辨的，其中的a、b、c均随机独立选取。与第2.2节中介绍的相同，IP-FE=(Setup,KeyDer,Encrypt,Decrypt)。

● Setup(l^{λ},l^{l})初始化整个系统，生成主公钥和主私钥，将全局安全参数λ和l作为输入参数，其中整数$s_1,\cdots,s_{\varsigma}\in Z_p$，最后返回密钥对(mpk,msk)。

$$(\mathrm{mpk},\mathrm{msk})\leftarrow \mathrm{Setup}(l^{\lambda},l^{l}) \tag{2}$$

$$\mathrm{msk}=s=(s_1,\cdots,s_{\varsigma})\leftarrow Z_p^{l} \tag{3}$$

● KeyDer(msk,$\boldsymbol{y}$)为数据使用者提供的向量$\boldsymbol{y}=(y_1,\cdots,y_l)$生成对应的sk，将sk作为获取$F(x)$的密钥。以主私钥以及$\boldsymbol{y}$为输入，返回密钥$\mathrm{sk}_y=<\boldsymbol{y},s>$。

● Encrypt(mpk,$\boldsymbol{x}$)使用主公钥对数据$\boldsymbol{x}=(x_1,\cdots,x_l)$，$x_i\in Z_p^{l}$进行加密，保护原始数据的隐私安全。将主公钥及数据$\boldsymbol{x}$作为输入，返回密文Ct。

$$\mathrm{Ct}=\left(\mathrm{Ct}_0,(\mathrm{Ct}_i)_{i\in[l]}\right) \tag{4}$$

$$\mathrm{Ct}_i=h_i^{r}\cdot g^{x_i},i\in[l] \tag{5}$$

其中，$Ct_0=g^r$，g是G的生成元，h、r是随机变量。

● Decrypt(mpk,Ct,sk_y)以主公钥mpk、密文Ct、密钥sk_y为输入，返回以g为基的离散对数。因此，由上述3个算法可以得出。

$$\text{Decrypt}(\text{mpk},\text{Ct},\text{sk})=\frac{\prod_{i\in[l]}\text{Ct}_i^{y^i}}{\text{Ct}_0^{\text{sk}_y}}=\frac{\prod_{i\in[l]}\left(g^{s_i r+x_i}\right)^{y^i}}{g^{r\left(\sum_{i\in[l]}y^i s^i\right)}}= \tag{6}$$

$$g^{\sum_{i\in[l]}y^i s^i r+\sum_{i\in[l]}y^i x^i-r\left(\sum_{i\in[l]}y^i s^i\right)}=g^{\sum_{i\in[l]}y^i x^i}=g^{<x,y>}$$

4.2 零知识证明

本文模型采用与Zcash、Filecoin等相同的零知识证明技术zk-SNARK，基于Groth16算法[20]构造。证明者需要证明自己知道一个秘密witness($a_{l+1},\cdots,a_m$)与statement($a_1\cdots,a_l$)，a_0=1能够满足如下计算式。

$$\sum_{i=0}^{m}a_i u_i(X)\cdot\sum_{i=0}^{m}a_i v_i(X)=\sum_{i=0}^{m}a_i w_i(X)+h(X)t(X) \tag{7}$$

其中，$u_i(X)$、$v_i(X)$、$w_i(X)$、$h(X)$、$t(X)$均与待证明的问题本身相关。

本文模型需要证明sk的有效性，即sk是通过正确的msk计算得到的，然后将该问题转化为相应的电路。由于zk-SNARK不能直接解决任何实际问题，需要将问题转化为多项式，并对多项式施加一系列约束才能进行后续的证明。随后，将电路转化为R1CS（rank-1 constraint system），并通过拉格朗日插值法转化为QAP(quadratic arithmetic program)。使用Groth16算法生成证明。具体来说，本文模型使用zk-SNARK生成证明的过程主要分为以下3步。

● Trust Setup生成CRS作为公共安全参数。首先选取$\alpha,\beta,\gamma,\delta,x\leftarrow Z_p^*$，$\tau=(\alpha,\beta,\gamma,\delta,x)$，然后计算$\sigma=([\sigma_1]_1,[\sigma_2]_2)$，其中$[\sigma_1]_1$为椭圆曲线$G_1$上的元素，$[\sigma_2]_2$为椭圆曲线$G_2$上的元素。具体的计算式如下。

$$\sigma_1=\left(\alpha,\beta,\delta,\left\{x^i\right\}_{i=0}^{n-1},\left\{\frac{\beta u_i(x)+\alpha v_i(x)+w_i(x)}{\gamma}\right\}_{i=0}^{l},\left\{\frac{\beta u_i(x)+\alpha v_i(x)+w_i(x)}{\delta}\right\}_{i=l+1}^{m},\left\{\frac{x^i t(x)}{\delta}\right\}_{i=0}^{n-2}\right) \tag{8}$$

$$\sigma_2=\left(\beta,\gamma,\delta,\left\{x^i\right\}_{i=0}^{n-1}\right) \tag{9}$$

值得注意的是，计算完σ后，原始的σ_1、σ_2通常被称为“有毒废料”，应该被销毁，否则整个零知识证明将不再安全。

● 选取$r,s\leftarrow Z_p$，选取与msk和$\boldsymbol{y}$长度相等的随机变量g、h。通过向量乘法将msk和$\boldsymbol{y}$隐藏，具体计算式如下。

$$c_1=g\cdot\boldsymbol{y} \tag{10}$$

$$c_2=g\cdot\boldsymbol{y}+h \tag{11}$$

$$v=\text{msk}\cdot g \tag{12}$$

将c_1、c_2、v作为计算证明的公开部分，r、s、$\boldsymbol{y}$作为生成证明的秘密部分。

● 计算证明$\pi=([A]_1,[C]_1,[B]_2)$，其中A、B、C的生成方式如下。

$$A=\alpha+\sum_{i=0}^{m}a_i u_i(x)+r\delta \tag{13}$$

$$B=\beta+\sum_{i=0}^{m}a_i v_i(x)+s\delta \tag{14}$$

$$C=\frac{\sum_{l+1}^{m}a_i\left(\beta u_i(x)+\alpha v_i(x)+w_i(x)+h(x)t(x)\right)}{\delta}+As+Br-rs\delta \tag{15}$$

最后将证明π传入智能合约进行验证。

$$[A]_1\cdot[B]_2=[\alpha]_1\cdot[\beta]_2+\sum_{i=0}^{l}a_i\left[\frac{\beta u_i(x)+\alpha v_i(x)+w_i(x)}{\gamma}\right]_1\cdot[\gamma]_2+[C]_1\cdot[\delta]_2 \tag{16}$$

5 安全性分析与实验

5.1 安全性分析

本文模型基于Abdalla M等人[21]提出的函数加密方案，引入区块链及零知识证明技术，实现隐私保护和数据安全共享。下面对本文模型进行安全性证明。

CSP被假定为诚实的，即CSP会执行用户的命令，但同时对用户的数据好奇。在本文模型中，原始数据首先经过用户加密后再上传存储，且密钥仅在TA处存放。因此数据的隐私安全得以保证。此外，为了防止CSP提供错误的加密数据给DU，在上传数据时，对数据进行哈希计算并对结果进行数字签名，随后将哈希值与数字签名一并存储在区块链上。因此，DU通过运行哈希算法，就可以轻易辨别数据是否被篡改，数据完整性也能得到保证。对于数据的可用性，用户可以选择将数据存储在多个CSP上，通过冗余的方式保证数据的可用性，但冗余产生的费用需要用户自行承担。

DO通过限制数据使用者对加密数据的处理方式，实现对数据使用拥有完全的控制权。DU支付费用后，由TA负责生成密钥和密钥的有效性零知识证明，保证DU除函数处理结果外不能获取任何信息。零知识证明在链上被验证后，智能合约自动完成支付。从整个网络来看，此次数据交易已经完成，DU可以随时取回密钥，解密获取结果，因此，本文模型是安全有效的。

5.2 实验结果及分析

本文实验利用以太坊搭建了本地5节点私有网络，节点宿主机均为VMware虚拟机，版本为Ubuntu 18.04，其配置为Intel Core i7 CPU，内存为32 GB。为了方便实验，选用PoA单节点打包区块（即出块），其余参数，如区块大小、gasLimit与Rinkeby Ethereum testnet保持一致。利用Solidity语言编写智能合约Market，实现DO链上数据发布、DU链上数据购买、零知识证明链上验证等。使用Rust语言实现链下计算复杂度高、计算密集的、不适合在区块链上进行的算法部分，如函数加密、零知识证明证据的生成。链下数据通过交易的形式传递至链上，链上交易通过智能合约中定义的交易事件，即对应的动作会触发预先设计的事件，通知链下CSP和其他相关方。本实验采用本地节点模拟CSP，节点配置同上述节点宿主机，网络带宽为1 000 MB。

在实验初始化阶段，等网络出块稳定后，通过Remix将Market部署在本地私有网络中，零知识证明采用ZoKrates组件，并对部署合约和各个函数消耗的Gas和时间进行测量，结果见**表1**。对比ChooseCt与UploadFn函数可以看出，如果DU使用DO预先设定的函数，其Gas消耗量仅为上传自定义函数的5%，上传自定义函数消耗的Gas随所求变量大小的变化如**图2**所示。使用自定义函数的Gas较高的主要原因在于，随着所求变量大小的增大，调用UploadFn函数时传入参数的大小也会增加。该函数会将参数

表1 部署合约和各个函数消耗的 Gas 与时间

名称	Gas	时间/s
Market(Deployed)	2 031 046	0.42
SetFnAndPrice	40 091	0.01
ChooseCt	49 210	0.01
UploadFn(para_size=3.8 KB)	930 930	0.22
VerifyZk	1 014 800	0.28

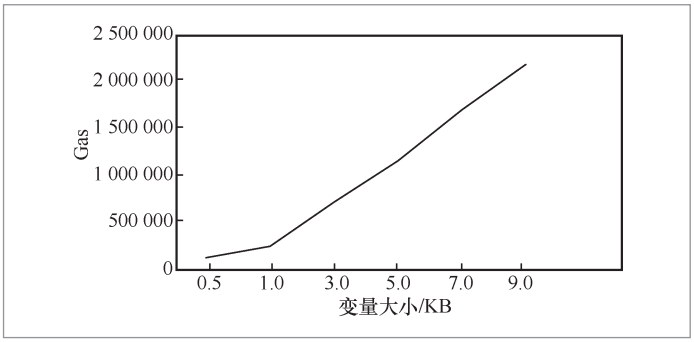

图 2 上传自定义计算函数消耗的 Gas 随所求变量大小的变化

写到链上并通过交易事件通知云服务商，而在以太坊中，存储成本远高于计算成本，因此DU可以尽可能选择预先定义的函数，以降低使用数据的成本。

内积函数加密算法性能会受到所求向量长度的影响，从而影响整个数据交易的性能，因此本文对不同长度的向量进行了性能测试，实验环境与上述以太坊节点宿主机相同，得到的实验结果如**图3**所示。由**图3**可以看出，随着向量长度的增加，加密和解密所需时间增长较块，但是派生密钥所需时间依然维持在较低水平，这也侧面证明了本文模型将加密和解密放到链下执行的正确性。

6 结束语

本文提出的基于区块链与函数加密的隐私数据安全共享模型，利用区块链及函数加密实现了数据的安全共享，借助函数加密的特性，用户可以自主控制数据使用者从密文中获取的信息量，避免使用传统公钥加密时的共享数据量不可控问题。并且，通过零知识证明生成相关计算的可信证明，实现了链上自动支付。本文模型在保证隐私数据的前提下，实现了“数据可用不可见”。

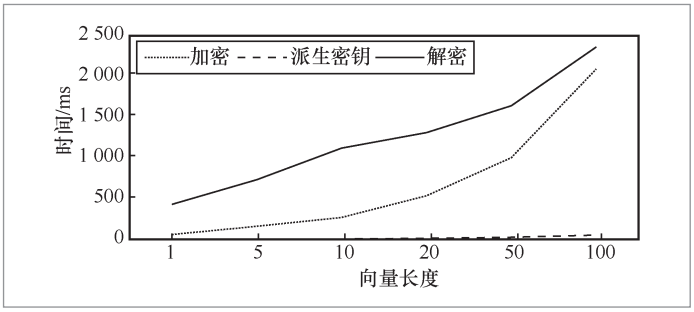

图 3 函数加密操作消耗的时间随向量长度的变化

参考文献:

[1] CHEN J C, XUE Y Z. Bootstrapping a blockchain based ecosystem for big data exchange[C]//Proceedings of 2017 IEEE International Congress on Big Data (BigData Congress). Piscataway: IEEE Press, 2017: 460-463.

[2] LIANG X P, ZHAO J, SHETTY S, et al. Integrating blockchain for data sharing and collaboration in mobile healthcare applications[C]//Proceedings of 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications. Piscataway: IEEE Press, 2017: 1-5.

[3] THEODOULI A, ARAKLIOTIS S, MOSCHOU K, et al. On the design of a blockchain-based system to facilitate healthcare data sharing[C]//Proceedings of 2018 17th IEEE International Conference on Trust, Security and Privacy In Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering. Piscataway: IEEE Press, 2018: 1374-1379.

[4] GORDON W J, CATALINI C. Blockchain technology for healthcare: facilitating the transition to patient-driven interoperability[J]. Computational and Structural Biotechnology Journal, 2018, 16: 224-230.

[5] AZARIA A, EKBLAW A, VIEIRA T, et al. MedRec: using blockchain for medical data access and permission management[C]// Proceedings of 2016 2nd International Conference on Open and Big Data. Piscataway: IEEE Press, 2016: 25-30.

[6] YANG H, YANG B. A blockchain-based approach to the secure sharing of healthcare data[C]//Proceedings of the Norwegian Information Security Conference. [S.l.:s.n.], 2017: 100-111.

[7] CAO S, ZHANG G X, LIU P F, et al. Cloud-assisted secure eHealth systems for tamper-proofing EHR via blockchain[J]. Information Sciences, 2019, 485: 427-440.

[8] 刘彦松, 夏琦, 李柱, 等. 基于区块链的链上数据安全共享体系研究[J]. 大数据, 2020, 6(5): 92-105.
LIU Y S, XIA Q, LI Z, et al. Research on secure data sharing system based on blockchain[J]. Big Data Research, 2020, 6(5): 92-105.

[9] YU K P, TAN L, ALOQAILY M, et al. Blockchain-enhanced data sharing with traceable and direct revocation in IIoT[J]. IEEE Transactions on Industrial Informatics, 2021, 17(11): 7669-7678.

[10] ZHANG Y, WEN J T. An IoT electric business model based on the protocol of bitcoin[C]//Proceedings of 2015 18th International Conference on Intelligence in Next Generation Networks. Piscataway: IEEE Press, 2015: 184-191.

[11] WÖRNER D, VON BOMHARD T. When your sensor earns money: exchanging data for cash with Bitcoin[C]//Proceedings of 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. New York: ACM Press, 2014: 295-298.

[12] ÖZYILMAZ K R, DOĞAN M, YURDAKUL A. IDMoB: IoT data marketplace on blockchain[C]//Proceedings of 2018 Crypto Valley Conference on Blockchain Technology. Piscataway: IEEE Press, 2018: 11-19.

[13] SHAFAGH H, BURKHALTER L, HITHNAWI A, et al. Towards blockchain-based auditable storage and sharing of IoT data[C]//Proceedings of 2017 on Cloud Computing Security Workshop. [S.l.:s.n.], 2017: 45-50.

[14] PAN J L, WANG J Y, HESTER A, et al. EdgeChain: an edge-IoT framework and prototype based on blockchain and smart contracts[J]. IEEE Internet of Things Journal, 2019, 6(3): 4719-4732.

[15] 张召, 田继鑫, 金澈清. 链上存证、链下传输的可信数据共享平台[J]. 大数据, 2020, 6(5): 106-117.
ZHANG Z, TIAN J X, JIN C Q. On-chain witness and off-chain transmission trustworthy data sharing platform[J]. Big Data Research, 2020, 6(5): 106-117.

[16] 袁勇, 王飞跃. 区块链技术发展现状与展望[J]. 自动化学报, 2016, 42(4): 481-494.
YUAN Y, WANG F Y. Blockchain: the state of the art and future trends[J]. Acta Automatica Sinica, 2016, 42(4): 481-494.

[17] SHAMIR A. Identity-based cryptosystems and signature schemes[C]//Proceedings of 1984 Workshop on the Theory and Application of Cryptographic Techniques. Heidelberg: Springer, 1985: 47-53.

[18] GOYAL V, PANDEY O, SAHAI A, et al. Attribute-based encryption for fine-grained access control of encrypted data[C]//Proceedings of the 13th ACM Conference on Computer and Communications Security. New York: ACM Press, 2006: 89-98.

[19] BONEH D, SAHAI A, WATERS B. Functional encryption: definitions and challenges[C]//Proceedings of 2011 Theory of Cryptography Conference. Heidelberg: Springer, 2011: 253-273.

[20] GROTH J. On the size of pairing-based non-interactive arguments[C]//Proceedings of 2016 Annual International Conference on the Theory and Applications of

Cryptographic Techniques. Heidelberg: Springer, 2016: 305–326.

[21] ABDALLA M, BOURSE F, CARO A D, et al. Simple functional encryption schemes for inner products[C]//Proceedings of the IACR International Workshop on Public Key Cryptography. Heidelberg: Springer, 2015: 733–751.

作者简介

李懿（1997– ），男，天津理工大学计算机科学与工程学院博士生，主要研究方向为区块链安全及应用、联邦学习。

王劲松（1970– ），男，博士，天津理工大学计算机科学与工程学院教授，中国计算机学会（CCF）理事，计算机病毒防治技术国家工程实验室副主任，主要研究方向为网络安全、数据智能、区块链等。

张洪玮（1990– ），男，博士，天津理工大学计算机科学与工程学院讲师，CCF区块链专业委员会执行委员，主要研究方向为区块链、数据安全、隐私保护。

收稿日期： 2022-04-01

通信作者： 王劲松，jswang@tjut.edu.cn

基金项目： 天津市研究生科研创新项目（No.2020YJSS067）；天津新一代人工智能重大专项（No.19ZXZNGX00080）

Foundation Items: Tianjin Graduate Science and Technology Innovation Project (No.2020YJSS067), The New Generation Artificial Intelligence Technology Major Project of Tianjin (No.19ZXZNGX00080)