·热点与综述·　　文章编号：1000-3428(2022)10-0013-08　　文献标志码：A　　中图分类号：TP309.2

区块链数据保密查询的不经意传输协议

刘　新[1]，胡翔瑜[1]，徐　刚[2]，陈秀波[3]

（1.内蒙古科技大学 信息工程学院，内蒙古 包头 014010；2.北方工业大学 信息学院，北京 100144；
3.北京邮电大学 网络与交换技术国家重点实验室，北京 100876）

摘　要：在区块链数据存储与查询过程中，由于区块链的透明性和公开性，全网所有用户均有可能获取查询者的数据信息，存在泄漏查询者隐私数据的风险。采用区块链链上-链下存储思想，设计区块链数据存储模型，引入代理重加密机制，将存储者加密后的数据分布式存储在链下，将存储者发送的索引信息和Merkle树根哈希值存储在链上，确保了数据的完整性、可靠性和可验证性，并减少了区块链数据对存储资源的占用。利用椭圆曲线加密算法设计区块链数据保密查询的不经意传输协议，使得全网所有用户均无法获取查询者的数据信息，保护了区块链数据传输过程中查询者的隐私。分析结果表明，该协议中查询者完成一次区块链上的不经意传输仅需$2n+2k+2$次椭圆曲线乘法运算，相比于现有不经意传输协议具有存储空间小、计算复杂度低等优势，并且在相同长度的密钥下具有更高的安全性。

关键词：区块链；保密查询；链上-链下存储模型；不经意传输；椭圆曲线加密算法

开放科学（资源服务）标志码（OSID）：

中文引用格式：刘新，胡翔瑜，徐刚，等. 区块链数据保密查询的不经意传输协议[J]. 计算机工程，2022，48(10)：13-20.

英文引用格式：LIU X，HU X Y，XU G，et al. Oblivious transfer protocol for confidentiality query of blockchain data[J]. Computer Engineering，2022，48(10)：13-20.

Oblivious Transfer Protocol for Confidentiality Query of Blockchain Data

LIU Xin[1]，HU Xiangyu[1]，XU Gang[2]，CHEN Xiubo[3]

（1.School of Information Engineering，Inner Mongolia University of Science and Technology，Baotou，Inner Mongolia 014010，China；2.College of Information，North China University of Technology，Beijing 100144，China；3.State Key Laboratory of Network and Switching Technology，Beijing University of Posts and Telecommunications，Beijing 100876，China）

【Abstract】During the data storage and query of a blockchain，owing to the transparency and openness of the blockchain，all users of a network may obtain information regarding the inquirer；thus，confidential information regarding the inquirer may be exposed.This study adopts the idea of a blockchain on-chain off-chain storage，designs a blockchain data storage model，introduces a proxy re-encryption mechanism for storing the encrypted data of the storer off-chain in a distributed manner，and stores the index information sent by the storer as well as the hash value of the root of the Merkle tree on the chain；this，in turn，ensures the integrity，reliability，and verifiability of data and also reduces the utilization of storage resources by the blockchain data.The elliptic curve encryption algorithm is used to design an Oblivious Transfer（OT）protocol for the confidential query of blockchain data，which prevents all the users in the entire network from obtaining the inquirer's information；this protects the privacy of the inquirer during data transmission.Analysis results reveal that the inquirer requires only $2n+2k+2$ elliptic curve multiplication operations to complete an OT on the blockchain when using this protocol.Compared with existing protocols，this OT protocol requires a smaller storage space，features lower computational complexity，and offers higher security for the same key length.

【Key words】blockchain；confidentiality query；on-chain off-chain storage model；Oblivious Transfer（OT）；elliptic curve encryption algorithm

DOI：10. 19678/j. issn. 1000-3428. 0063861

0　概述

随着区块链技术的快速发展，人们对于区块链隐私保护问题越来越重视，如何在区块链上存储数据并保护数据隐私成为热点研究方向[1-2]。现有研究主要集中于实现区块链交易双方身份和交易内容

基金项目：国家自然科学基金（92046001）；内蒙古自治区自然科学基金（2021MS06006）；内蒙古自治区科技重大专项（2019ZD025）；内蒙古自治区纪检监察大数据实验室开放项目（IMDBD2020020）；包头市科技计划项目（YF2020013）；北京市教委基本科研业务费资助项目（110052972027）；北方工业大学科研启动基金（110051360002）。

作者简介：刘　新（1983—），男，副教授、博士，主研方向为信息安全、区块链；胡翔瑜，硕士研究生；徐　刚，博士；陈秀波，教授、博士。

收稿日期：2022-01-27　　**修回日期：**2022-03-04　　**E-mail：**gx@ncut.edu.cn

的隐私保护[3-5]，用户将数据存储在区块链上，在便于交易和计算的同时还需解决数据管理和存储空间问题。文献[6]提出一种基于区块链和代理重加密的数据共享方案，利用区块链中的处理节点作为代理服务器并加密数据，在确保数据安全的同时可抵抗合谋攻击。文献[7]设计一种链上-链下共同存储方式，将存储者的数据访问地址以加密的形式存储在链上，将用户的大量数据存储在链下，若查询者需要访问该用户数据，则需要被授予存储者访问令牌，并交予第三方获得真正的地址，显然该系统存在第三方泄漏数据地址和查询者信息的风险。针对数据共享时存在数据泄漏的情况，文献[8]提出一种可问责的数据共享方案(ADS)，该方案利用不经意传输(Oblivious Transfer，OT)和零知识证明，将接收方的私钥隐藏在共享数据中，若接收方泄漏共享数据，则发送方可以对其进行问责，从而获取接收方的私钥，但该方案存在发送方可能篡改数据的问题。文献[9]建立一种基于链上-链下相结合的日志安全存储与检索模型，利用区块链分布式存储技术，保证了数据的机密性和完整性，并通过链上索引的方式进行数据检索，然而该方案没有解决用户在检索与传输数据时导致个人信息泄漏的问题。文献[10]提出一种将医疗数据存储在区块链上的方案，并对数据共享、存储、访问和计算进行隐私保护方面的评估和分析，同时构建一个模块化的混合隐私保护框架，该框架基于保护患者的隐私信息来管理不同类型的医疗数据，对电子病例、消费者基因数据等信息进行保密，但攻击者可以通过收集用户共享、传输的医疗数据推测出用户的隐私敏感信息。

在区块链数据存储和数据查询过程中，当查询者获取数据时，全网每个节点均有可能获取查询者的敏感信息，因此对于查询者的隐私保护显得尤为重要。文献[11]提出一种公平的大数据交换方案(FAPS)，采用RSA算法与AES算法相结合的方式进行不经意传输，保护了交易双方的数据隐私，且不需要可信第三方参与传输，但该方案存在交易一方恶意篡改数据的情况。文献[12]设计一种隐私保护下的区块链关键词搜索方案，使得数据提供者无法了解查询者搜索内容的相关信息，同时利用了区块链可验证数据的性质，使得恶意方无法篡改数据。在该方案中，数据存储者在区块链上加密存储数据，各个节点相互验证后按顺序将数据存储在区块链上，导致了区块链存储量大、节点计算复杂、传输效率低下及节点之间验证存在延迟等问题。文献[13]提出一种基于不经意传输和区块链的隐私保护数据传输方案，并将其应用于智能医疗中，保护了医生和病人的数据隐私，同时采用代理重加密技术实现密文之间的转换，但依然存在海量数据分布式存储在区块链上导致的存储量大、资源浪费等问题。

本文采用链上-链下数据存储方式，引入代理重加密机制[14-15]，将存储者加密后的数据采用MapReduce进行归类后分布式存储在链下数据存储层，并将存储者发送的索引信息存储在链上的块身中，使得查询者能根据索引信息找到需要的数据，不论是存储者本身还是其他用户均无法随意篡改数据。查询者和存储者利用椭圆曲线加密算法[16-18]进行不经意传输，使得所有用户均无法获取查询者数据信息，同时存储者的私钥和原始数据也不会被泄漏。

1 基础知识

1.1 不经意传输协议

不经意传输协议[19]是密码学的基本工具，在安全多方计算中具有广泛应用[20]。该协议保证接收方获取发送方的某个数据后，发送方不知道接收方获取了具体哪个数据，同时接收方无法获取发送方的其他数据[21-22]，目前研究最多的是OT_n^k协议[23-25]，其中，n为发送方的所有数据个数，k为接收方获取的数据个数。

1.2 Merkle树数据结构

Merkle树数据结构[26-27]主要用来快速归纳和校验数据的完整性，将数据分组进行哈希运算，向上不断递归运算产生新的哈希节点，最终只剩下一个Merkle树根哈希值，如图1所示，其中$H(m)$表示对数据m取哈希值。Merkle树具有不可篡改、可验证、高效率等特点，运用在区块链中极大地提高了区块链运行效率和可扩展性。

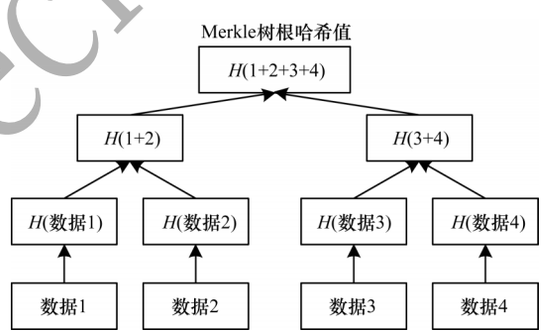

图1 Merkle树数据结构

Fig.1 Merkle tree data structure

1.3 区块链存储框架

由于数据在区块链上进行同步时会占用大量空间，因此本文根据链上-链下相结合的存储思想，设计链上存储索引信息、链下存储原始数据的存储框架，释放了区块链上的大量存储空间。将区块链分为网络层和数据存储层，链上的网络层全网公示，用于查询者选择区块链数据，链下的数据存储层存储大量的加密信息，如图2所示。

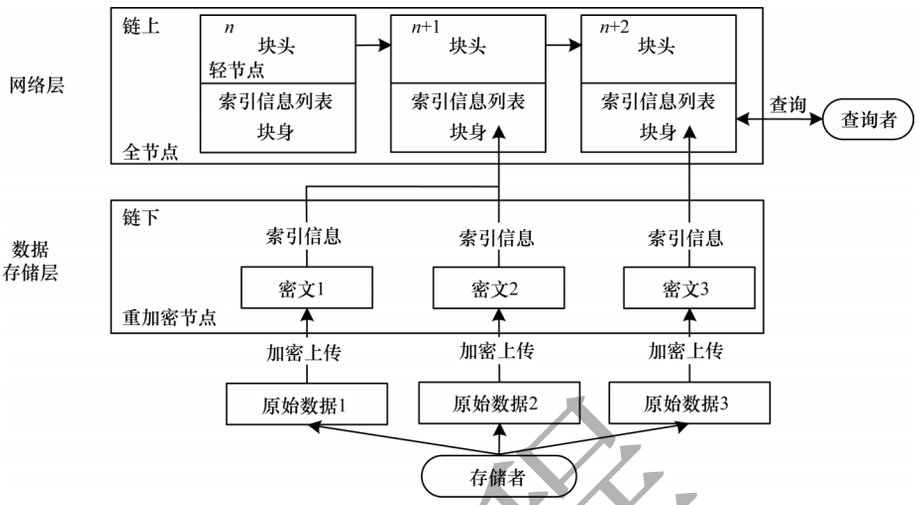

图2　区块链存储框架

Fig.2　Blockchain storage framework

在图2中,索引信息由存储者生成,其中包括数据标题、密文哈希值等,存储者将密文发送至重加密节点存放在链下数据存储层,将索引信息发送至全节点存放在链上网络层的块身中形成索引信息列表。该框架中有轻节点、全节点、重加密节点3类,其中:轻节点负责记录区块链的块头以及促进区块链的运转;全节点负责记录整个链上的信息,包含块头和块身,为查询者提供块身中的索引信息;重加密节点负责保存链下数据存储层的大量数据,而且诚实的重加密节点数量大于全部重加密节点数量的1/2,即可保证数据不可被篡改。

1.4　基于MapReduce的区块链数据存储模型

MapReduce是一种编程模型,用于大规模数据集的并行运算,从而提高计算效率。重加密节点收到发布的新区块后就进行一次MapReduce运算,将新区块中所有存储者的数据按照存储者用户名进行分类,存储在数据存储层,以便在完成不经意传输时节省筛选存储者数据的时间。基于MapReduce的区块链数据存储框架如图3所示。

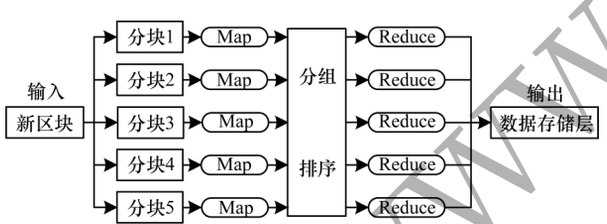

图3　基于MapReduce的区块链数据存储框架

Fig.3　MapReduce-based blockchain data storage framework

在MapReduce执行过程中,重加密节点在产生各个分节点的同时进行并行运算,首先在Map阶段将数据进行切割,然后在Shuffle阶段进行分组和排序,接着在Reduce阶段进行归约,最后重加密节点将所有的数据按照存储者用户名分类存储在数据存储层中。重加密节点通过分布式存储的方式将数据备份在各个分节点处,分节点将密文取哈希后与链上索引信息中的密文哈希值进行对比验证,保证了数据的可用性和完整性。

2　区块链数据保密查询的OT协议

本文设计的区块链数据保密查询的不经意传输协议分为数据存储和不经意传输两部分。在数据存储部分:重加密节点将存储者加密后的数据采用MapReduce进行归类后,分布式存储在链下数据存储层,确保了数据的完整性;全节点将存储者发送的索引信息存储在链上的块身中,使得查询者能根据索引信息找到需要的数据,不论是存储者本身还是各个节点,均无法随意篡改数据,确保了数据的可靠性和可验证性。在不经意传输部分,查询者和存储者通过重加密节点进行不经意传输协议后,所有节点和存储者无法得知查询者获取了哪个数据,同时存储者的私钥和原始数据也不会被泄漏。

在区块链中,诚实的重加密节点数量大于全部重加密节点数量的1/2,即可保证协议能够顺利执行。此外,在区块链上存储的原始数据均视为正确数据,不考虑输入数据为错误的情况。

2.1　链上-链下数据存储

链上网络层中的每个区块分为块头和块身,块头由轻节点进行分布式存储,上一区块头的哈希值记录在本块头中,防止数据被篡改,块头中还记录了随机数、Merkle树根哈希值、时间戳等信息。索引信息存放在块身中形成索引信息列表,按照从左到右的顺序排列,由全节点进行记录,使得查询者根据块身中的索引信息列表找到所需的数据。区块链网络层中的数据存储框架如图4所示。

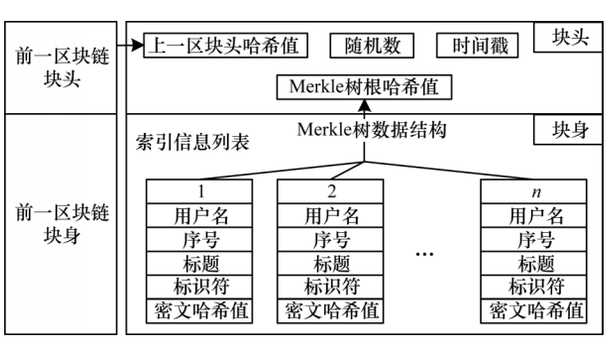

图4　区块链网络层中的数据存储框架

Fig.4　Data storage framework in the blockchain network layer

存储者利用椭圆曲线加密算法对上传的信息进行加密得到密文和标识符，再用SHA256哈希函数将密文转为哈希值，同时为该密文选取标题和序号。存储者将密文发送至重加密节点存储在数据存储层中，再将自己的用户名和该密文的序号、标题、标识符、密文哈希值组成一个索引信息发送至全节点。序号表示执行不经意传输协议时索引信息的发送顺序，标题为查询者提供信息类别，标识符用于加密查询者公钥，密文哈希值用于数据不经意传输完成后的验证。全节点将所有索引信息公开存储在链上块身中形成索引信息列表，同时用SHA256哈希函数依次将每一条索引信息转化为哈希值，用Merkle树数据结构的方式形成一个根哈希值存储在块头中。

2.2 不经意传输过程

重加密节点通过对密文进行二次加密，使得存储者不需要提供私钥，查询者仅利用自己的私钥就能完成对二次加密后的密文进行解密，交互过程具体如下：

1）查询者确定获取数据的序号后，利用索引信息中的标识符加密自己的公钥发送给存储者。

2）存储者将自己的私钥与查询者发送的加密数据进行结合，产生重加密密钥，并将重加密密钥发送给重加密节点。

3）重加密节点在数据存储层中收到存储者上传的密文，利用重加密密钥依次对其进行二次加密，然后依次发送给查询者。

4）查询者按序号找到重加密节点发送的二次加密密文，利用自己的私钥进行解密获取存储者的原始数据。

区块链不经意传输过程如图5所示。

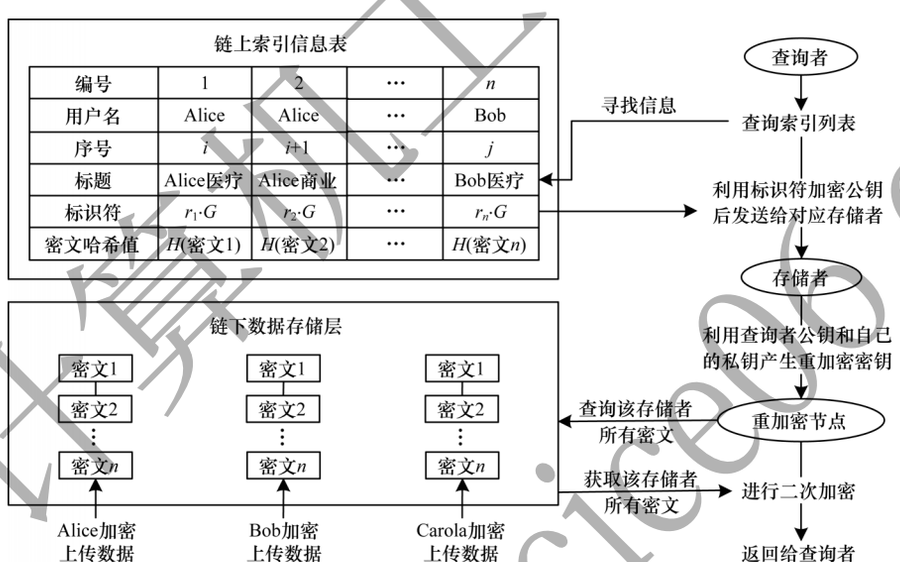

图5 区块链不经意传输过程

Fig.5 Process of blockchain oblivious transfer

2.3 不经意传输协议构建与安全性证明

2.3.1 协议构建

假设Alice为数据存储者，Bob为查询者，Bob想要获取Alice存储在区块链上的某个数据，但不想让全网和Alice知道自己获取了哪个数据。协议分为数据存储阶段和不经意传输阶段，Alice在区块链正常运行下执行数据存储阶段，Bob在需要获取区块链上的数据时执行不经意传输阶段。在整个协议过程中，均采用椭圆曲线加密系统进行加解密。协议中的符号含义如表1所示，其中$i\in(1,2,\cdots,n)$。

假设用户Alice拥有原始数据$m_1,m_2,\cdots,m_n$，公钥p_A和私钥s_A，随机数r_i。Bob拥有公钥p_B和私钥s_B。Alice利用私钥和随机数将数据存储到区块链中，Bob需要查询数据时发起申请，在通过身份验证后，开始与Alice进行交互，协议算法伪代码如算法1所示。

表1 协议中的符号含义

Table 1 The meaning of symbols in the protocols

符号	含义
p_A和p_B	Alice和Bob的公钥
s_A和s_B	Alice和Bob的私钥
m_i	原始数据
r_i	加密随机数
C_i	链下密文
C_i'	标识符
X	加密数据
P_i	解密数据
x	Bob选择数据序号
M_i	原始数据编码点
W_i	二次加密密文
K	重加密密钥

算法1 区块链保密数据查询的不经意传输算法

输入 Alice的数据m_i,Alice的公私钥p_A和s_A,Bob的公私钥p_B和s_B,椭圆曲线编码Code,椭圆曲线加密算法E,解密算法D

输出 明文信息

1.$m_i=m_1,m_2,\cdots,m_n$;

2.$Code(m_i)=M_1,M_2,\cdots,M_n$;

3.$E_{s_A}(m_i)=C_1,C_2,\cdots,C_n$;

4.$E_{r_i}(m_i)=C'_1,C'_2,\cdots,C'_n$;

5.$E_{p_B}(C'_x)=C'_x+p_B=X$;

6.$E_{s_A}(X)=s_A\cdot X=K$;

7.$E_K(C_i)=K-C_i=W_i$;

8.$D_{s_B}(W_x)=s_B\cdot p_A-W_x=M_x$

区块链保密数据查询的不经意传输协议包括数据加密和数据保密查询两个阶段,具体如下:

1)数据加密阶段:Alice选择s_A作为私钥,通过椭圆曲线加密算法生成公钥$p_A=s_A\cdot G$,将原始数据编码成椭圆曲线上的点后进行公钥加密得到密文C_i,每加密一个原始数据时Alice需要选择一个随机数r_i,同时利用这个随机数生成一个标识符C'_i,加密过程如下:

$$C=\begin{cases}C_i=M_i+r_i\cdot p_A\\C'_i=r_i\cdot G\end{cases}\tag{1}$$

2)数据保密查询阶段:(1)Bob选择随机数s_B作为私钥,产生公钥$p_B=s_B\cdot G$,利用索引信息中的标识符和序号x计算$X=C'_i+p_B$,然后将X发送给Alice;(2)Alice计算$K=s_A\cdot X$得到重加密密钥,并将K发送给重加密节点;(3)重加密节点在链下存储层中找到Alice所有的密文C_i,利用重加密密钥K依次计算$W_i=K-C_i$,并将W_i返回给Bob;(4)Bob通过序号找到对应的W_x(即i取x),利用私钥s_B和Alice的公钥p_A计算$M_x=s_B\cdot p_A-W_x$,并得到Alice原始数据。至此协议结束。

2.3.2 正确性分析

正确性分析步骤具体如下:

1)假设在数据保密查询阶段中第4步解密结果是正确的,Bob对W_x进行如下解密过程:

$$\begin{aligned}M_x=s_B\cdot p_A-W_x=s_B\cdot s_A\cdot G-K+C_x=\\s_B\cdot s_A\cdot G-s_A\cdot C'_x-s_A\cdot p_B+C_x=\\s_B\cdot s_A\cdot G-s_A\cdot r_x\cdot G-s_A\cdot s_B\cdot G+\\M_x+r_x\cdot s_A\cdot G=M_x\end{aligned}\tag{2}$$

2)Bob无法解密Alice的其他数据,假设Bob选择第y个二次加密密文W_y进行解密,由于$y\neq x$,因此$r_y\neq r_x$,解密过程如下:

$$\begin{aligned}s_B\cdot p_A-W_y=s_B\cdot s_A\cdot G-s_A\cdot p_B-\\s_A\cdot r_x\cdot G+M_y+r_y\cdot s_A\cdot G=\\r_y\cdot s_A\cdot G-s_A\cdot r_x\cdot G+M_y\neq M_y\end{aligned}\tag{3}$$

3)Bob和重加密节点无法获取Alice的私钥,因为已知$K=s_A(C'_x+p_B)$中的K、p_B和C'_i,无法推导出s_A的值,已知$W_i=K-C_i=s_A\cdot p_B+s_A\cdot r_i\cdot G-M_i-r_i\cdot s_A\cdot G=s_A\cdot s_B\cdot G-M_x$中的$W_i$、$K$、$p_B$、$r_i\cdot G$、$M_i$和$s_B\cdot G$,无法由任何一个等式推导出$s_A$的值。

4)Alice和重加密节点无法知道Bob获取了哪个数据,在第2步中,Alice收到Bob发送的加密数据X后,可利用标识符依次计算$P_i=X-C'_i$解密得到Bob的公钥,但得到n个数据时却不知道哪个是Bob的公钥,每个数据都有$1/n$的概率是Bob的公钥,因此Alice得到的数据均不可区分,具有很好的隐私性。

2.3.3 威胁模型分析

本文提出的区块链数据保密查询的不经意传输协议具有识别恶意节点的功能,根据安全性假设,通过以下2种情况进行分析:

情况1 若重加密节点篡改链下存储层的数据,查询者则无法获取正确结果。

查询者获取数据后可利用区块链的公开性来验证得到的数据是否正确,首先利用存储者的公钥加密原始数据后获取哈希值,然后与链上索引信息中的密文哈希值进行对比,若相等则说明验证成功,该重加密节点是诚实的;若不相等,则查询者可向区块链申请,让存储者将重加密密钥发送给另一位新的重加密节点,若新的重加密节点返回的二次加密密文能让查询者通过验证,则表明之前的重加密节点私自篡改了数据,并将其视为恶意节点。

情况2 若存储者发送错误的重加密密钥给重加密节点,查询者则无法获取正确结果。

当诚实的重加密节点数量大于1/2时,若查询者在多次验证失败向区块链发起申请后(即存储者多次发送重加密密钥给新的重加密节点后),验证新重加密节点返回的数据均失败,则表明存储者提供了错误的重加密密钥,并将其视为恶意节点。

在识别出恶意节点后,采用基于声誉信任机制[28]或基于激励机制[29]的管理方法来抵制恶意节点。P2P网络中最重要的特征之一就是需要节点积极参与,对提供正确数据的节点进行奖励,对恶意节点进行惩罚,节点通过不断赚取声誉来获取信任,从而获得更多的奖励。

2.3.4 安全性证明

本文使用模拟范例方法证明协议的安全性[30],构造两个模拟器S_1与S_2代替参与双方执行协议,如果能证明协议中任意一方无法获取其他参与者的信息,或者获取的数据是不可区分的,即可证明该协议是安全的。

定义1(安全性[30]) 对于协议π和函数$f(x,y)$,如果存在概率多项式时间算法使得式(4)、式(5)成立,则认为π保密计算了f,表明π具有较高的安全性。

$$\begin{aligned}&\{S_1(x,f_1(x,y)),f_2(x,y)\}\overset{c}{\equiv}\\&\{\text{view}_1^{\pi}(x,y),\text{output}_2^{\pi}(x,y)\}\end{aligned}\tag{4}$$

$$\begin{aligned}&\{f_1(x,y),S_2(y,f_2(x,y))\}\overset{c}{\equiv}\\&\{\text{output}_1^{\pi}(x,y),\text{view}_2^{\pi}(x,y)\}\end{aligned}\tag{5}$$

其中:Alice和Bob得到的信息序列分别记为$\text{view}_1^{\pi}(x,y)=(x,r^1,m_1^1,m_2^1,\cdots,m_i^1)$和$\text{view}_2^{\pi}(x,y)=(x,r^2,m_1^2,m_2^2,\cdots,m_i^2)$,输出结果分别记为$\text{output}_1^{\pi}(x,y)$和$\text{output}_2^{\pi}(x,y)$,$\overset{c}{\equiv}$表示计算上不可区分。若要证明一个两方计算协议是安全的,则必须构造满足式(4)和式(5)的模拟器S_1和S_2。

构造模拟器S_1和S_2分别代表Alice和Bob，通过模拟范例模拟Alice和Bob的执行过程，从而证明协议的安全性。在证明过程中，假设攻击者会对重加密节点进行攻击，从而获取重加密节点的信息，因此在view_1^{π}、view_2^{π}中会加入重加密节点的数据并证明其不可区分。由于S_1在执行时希望推导出S_2的公钥，因此在收到加密数据X后，会试图计算出$p_{x'}=X-C'_{x'}$与p_x进行对比，S_2执行时希望获取S_1的其他数据M。

定理1 区块链保密数据查询的不经意传输协议是安全的。

证明

在协议中，$\text{view}_1^{\pi}(m_i,x)=\{m_i,D(X),E(C),C\}$，其中，$m_i$为Alice原始数据，$x$为Bob选择的数据序号，$i\in(1,2,\cdots,n)$，$D(X)$是Alice解密Bob发送的数据，$E(C)$是Alice重加密节点二次加密密文。$\text{view}_2^{\pi}(m_i,x)=\{x,D(W),W,P(m_x,x)\}$，其中，$D(W)$是Bob解密返回数据，$P(m_x,x)$是Bob输出结果。$f_1(m_i,x)$为$S_1$执行协议的函数，$f_2(m_i,x)$为$S_2$执行协议的函数，$\text{output}_2^{\pi}(m_i,x)=P(m_x,x)$，输出结果只有Bob知道。

模拟器S_1执行过程如下：

1）根据$f_1(m_i,x)$，S_1选择x'，使得$f_1(m_i,x)=f_1(m_i,x')$。

2）S_1解密$D(X')$得到$p_{x'}$，其中$p_{x'}=X-C'_{x'}$。将$p_{x'}$与p_x进行对比，其中$p_x=X-C'_x$。由于$C'_x\overset{c}{\equiv}C'_{x'}$，因此$p_{x'}p_x$。

3）S_1加密计算$E(C_{x'})$，得到$K=s_A\cdot p_{x'}+s_A\cdot C'_{x'}$和$W_{x'}=K-C_{x'}$。

令$\{S_1(m_i)\}=\{m_i,D(X'),E(C_{x'}),C\}$。由于$p_{x'}\overset{c}{\equiv}p_B\rightarrow D(X')\overset{c}{\equiv}D(X),E(C_{x'})\overset{c}{\equiv}E(C),W_{x'}\overset{c}{\equiv}W_x$，因此式(6)成立：

$$\{S_1(m_i,f_1(m_i,x)),f_2(m_i,x)\}_{m_i,x}\overset{c}{\equiv}\{\text{view}_1^{\pi}(m_i,x),\text{output}_2^{\pi}(m_i,x)\}_{m_i,x}\tag{6}$$

可见，Alice在计算上不可区分，Bob是安全的。

模拟器S_2执行过程如下：

1）S_2选择序号为x的标识符C'_x加密自己的公钥p_B，得到$X=C'_x+p_B$。

2）根据$f_2(m_i,x)$，S_2选择x''，使得$f_2(m_i,x)=f_2(m_i,x'')$。

3）S_2利用私钥s_B和Alice公钥p_A进行$D(W_{x''})$解密第x''个数据$W_{x''}$得到M，此时$M=s_B\cdot p_A-W_{x''}=s_A\cdot r_{x''}\cdot G-s_A\cdot r_x\cdot G+M_{x''}\neq M_{x''}$，$M\neq M_i$。

令$\{S_2(x,P(m_x,x))\}=\{x,D(W_{x''}),W_{x''},P(m_{x''},x'')\}$。由于$D(W_{x''})\overset{c}{\equiv}D(W_i)$，$W_{x''}\overset{c}{\equiv}W_i$，且$P(m_i,x'')\neq P(m_i,x)$，因此式(7)成立：

$$\{f_1(m_i,x),S_2(x,f_2(m_i,x))\}_{m_i,x}\overset{c}{\equiv}\{\text{output}_2^{\pi}(m_i,x),\text{view}_2^{\pi}(m_i,x)\}_{m_i,x}\tag{7}$$

可见，Bob在计算上不可区分，Alice是安全的。

证明完毕。

3 性能分析

3.1 存储空间分析

本文采用的区块链数据存储结构分为链上网络层和链下存储层，分别对其计算效率和存储空间进行分析。假设存储者用户数为t，每个用户上传了n个长度为l的数据，椭圆曲线采用secp256k1曲线。

在链下存储层中，存储者利用椭圆曲线加密算法加密数据得到密文和标识符，计算效率为$2nt$次椭圆曲线乘法运算。在存储空间方面，重加密节点分布式存储所有存储者的密文，每个密文大小为512 bit，则重加密节点需要的存储空间大小为$512nt$ bit。

在链上网络层中，全节点共收集nt个索引信息，将其递归运算得到Merkle树根哈希值，计算效率为$2nt-1$次SHA256哈希运算。在存储空间方面，每个索引信息标识符为512 bit，密文哈希值为256 bit，用户名、序号、标题共计256 bit，则每一个块身索引信息列表大小为$1\,024nt$ bit。块头中轻节点仅存储Merkle树根哈希值和前一区块哈希值，共512 bit，因此块头存储大小可忽略不计。

3.2 计算效率分析

将本文提出的不经意传输协议与文献[11,13]和文献[21-24]协议从以下2个方面进行计算效率对比，对比结果如表2所示：

1）计算复杂度。计算复杂度是评价协议的重要因素，对比协议在传输过程中采用加减法运算、异或运算、哈希运算、模指数运算、椭圆曲线上的加法和乘法运算等，比较耗时的为模指数运算和椭圆曲线上的乘法运算，其中，$r=[l/L]$，l为明文的比特长度，L为映射到群G后元素的最长明文比特长度，n表示发送方的所有数据数量，k表示接收方获取的数据数量。文献[11]在初始化阶段需要$2n$次模指数运算，在不经意传输阶段需要$2kn+2k$次模指数运算，共计$(2k+2)n+2k$次模指数运算；文献[13]在准备和注册阶段需要$n+4$次模指数运算，在不经意传输阶段需要$2n+k+4$次模指数运算，共计$3n+k+6$次模指数运算。文献[21]由于每个接收方单独获取信息，因此发送方需要重新计算对称密钥，共计$kn+k+2$次模指数运算；文献[22]将大部分指数运算外包至2个云服务器从而降低发送方和接收方的运算效率，共计$2.5n+2k+3$次模指数运算；文献[23]仅需要一个云服务器，共计$3.25n+k+1$次模指数运算；文献[24]利用椭圆曲线加密算法进行不经意传输，其中方案1所设计的OT_n^k协议需要$rn+n+3k+1$次乘法运算，方案2通过增加接收方运算量的方式，从而减少整个协议的计算复杂度，依然需要$rn+3k$次乘法运算，其中r表示消息明文长度/映射到群G后最长明文长度。本文协议(一次接收k个数据)在数据存储阶段需要2次椭圆曲线乘法运算，在不经意传输阶段发送方加密密文需要$2n$次椭圆曲线乘法运算，计算重加密密钥需要k次椭圆曲线乘法运算，接收方解密时需要k次椭圆曲线乘法运算，共计$2n+2k+2$次椭圆曲线乘法运算。

2）通信复杂度。通信复杂度通常用交互的轮数来衡量。文献[11]协议需要3轮交互，文献[13]协议需要3轮交互，文献[21]协议需要2轮交互，文献[22-23]协议均需要3轮交互，文献[24]方案1需要3轮交互，方案2需要2轮交互。在本文协议中，存储者与查询者只需要2轮交互。

表2　计算效率对比

Table 2　Comparison of computational efficiency

协议	计算复杂度	通信复杂度
文献[11]协议	$3n+2k+6$次模指数运算	3轮
文献[13]协议	$(2k+2)n+2k$次模指数运算	3轮
文献[21]协议	$kn+k+2$次模指数运算	2轮
文献[22]协议	$2.5n+2k+3$次模指数运算	3轮
文献[23]协议	$3.25n+k+1$次模指数运算	3轮
文献[24]方案1	$(r+1)n+3k+1$次椭圆曲线乘法运算	3轮
文献[24]方案2	$rn+3k$次椭圆曲线乘法运算	2轮
本文协议	$2n+2k+2$次椭圆曲线乘法运算	2轮

通过表2可得出：在计算复杂度上，由于椭圆曲线上的乘法运算优于模指数运算，本文协议相比文献[11，13]和文献[21-23]在计算效率上有所提升，当明文长度l大于$2L$时，相比文献[24]中的2种方案，本文协议计算效率较高；在通信复杂度上，本文协议相比文献[11，13]、文献[22-23]和文献[24]中的方案1少1轮，与文献[21]和文献[24]中的方案2的通信交互轮数相同。

4　实际应用结果

应用1　目前，医疗大数据隐私保护方案能对患者的原始数据进行保密，但并没有保护查询者的隐私，当查询者在访问数据库后，管理员通过获取查询者的访问记录，可推测出查询者患有某种疾病。在区块链中，所有的节点均有可能得知查询者获取的数据信息，导致了泄漏查询者的个人隐私。如图6所示，利用本文设计的区块链保密查询的不经意传输协议，不仅能加密区块链数据库中的其他数据，而且能保护查询者的个人隐私。

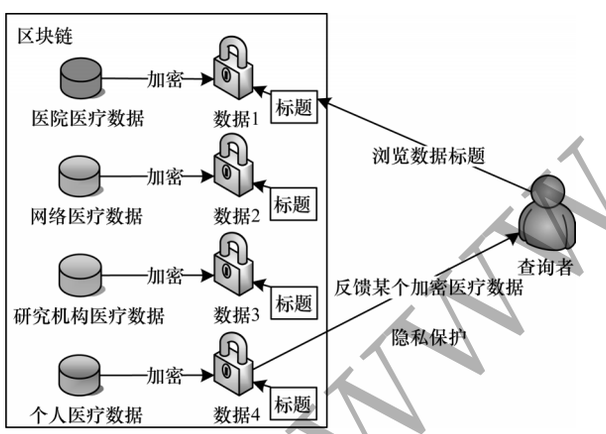

图6　本文协议在医疗领域中的应用

Fig.6　Application of the proposed protocol in the medical field

应用2　在大数据背景下，网上浏览信息或网购时的访问记录也可能暴露个人隐私。例如，网站下载的视频、交易商品的数据、查阅的资料等，即使这些数据的地址加密后存储在区块链上，当访问者在获取这些数据的下载地址后，通过分析访问者下载的数据，即可推断出其生活轨迹、个人喜好等。如图7所示，利用本文设计的区块链保密查询的不经意传输协议，能有效地保护访问者的个人隐私。

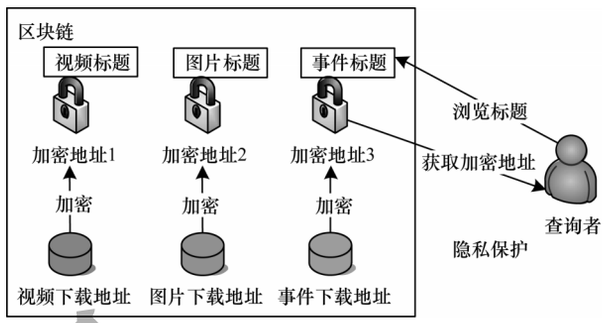

图7　本文协议在大数据领域中的应用

Fig.7　Application of the proposed protocol in the big data field

5　结束语

目前，防止隐私数据泄漏已成为区块链隐私保护的研究热点。对于数据的隐私保护不局限于原始数据本身，查询者的访问信息也需要加密，不经意传输协议为解决这类问题提供了思路。本文采用区块链链上-链下存储思想，引入代理重加密机制，设计区块链数据保密查询的不经意传输协议，保护了用户和查询者的隐私。分析结果表明，该协议计算效率高，占用存储空间少，保证了数据的完整性、可靠性和可验证性，同时能有效识别恶意节点，阻止敌手恶意行为。下一步将在区块链链上-链下数据存储模型下引入重加密节点，并结合零知识证明和同态加密等技术，设计能够抵抗恶意敌手攻击的隐私保护协议，在保护区块链用户隐私的前提下实现更高效安全的数据存储与传输。

参考文献

[1] CHENG L C，LIU J Q，JIN Y，et al. Account guarantee scheme：making anonymous accounts supervised in blockchain[J]. ACM Transactions on Internet Technology，2021，21(1)：1-11.

[2] 李莉，杜慧娜，李涛. 基于群签名与属性加密的区块链可监管隐私保护方案[J]. 计算机工程，2022，48(6)：132-138.
LI L，DU H N，LI T. Blockchain supervisable privacy protection scheme based on group signature and attribute encryption[J]. Computer Engineering，2022，48(6)：132-138. (in Chinese)

[3] LIU Q Y，LIU Z，LONG Y，et al. Making Monero hard-to-trace and more efficient[C]//Proceedings of the 18th IEEE International Conference on Trust，Security and Privacy in Computing and Communications/13th IEEE International Conference on Big Data Science and Engineering(TrustCom/BigDataSE). Washington D. C.，USA：IEEE Press，2019：514-521.

[4] ZHANG Y C，LONG Y，LIU Z，et al. Z-Channel：scalable and efficient scheme in Zerocash[M]. Berlin，Germany：Springer，2018.

[5] 李旭东，牛玉坤，魏凌波，等. 比特币隐私保护综述[J]. 密码学报，2019，6(2)：133-149.
LI X D，NIU Y K，WEI L B，et al. Overview on privacy

protection in bitcoin[J]. Journal of Cryptologic Research, 2019, 6(2): 133-149. (in Chinese)
[6] OBOUR A, KWAME O B, XIA Q, et al. A secured proxy-based data sharing module in IoT environments using blockchain[J]. Sensors (Basel, Switzerland), 2019, 19(5): 1235.
[7] 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, 2019, 15: 1746-1761.
[8] HUANG C, LIU D X, NI J B, et al. Achieving accountable and efficient data sharing in industrial Internet of Things[J]. IEEE Transactions on Industrial Informatics, 2021, 17(2): 1416-1427.
[9] 吕建富,赖英旭,刘静. 基于链上链下相结合的日志安全存储与检索[J]. 计算机科学, 2020, 47(3): 298-303.
LÜ J F, LAI Y X, LIU J. Log security storage and retrieval based on combination of on-chain and off-chain[J]. Computer Science, 2020, 47(3): 298-303. (in Chinese)
[10] MIYACHI K, MACKEY T K. hOCBS: a privacy-preserving blockchain framework for healthcare data leveraging an on-chain and off-chain system design [J]. Information Processing & Management, 2021, 58(3): 102535.
[11] LI T T, REN W, XIANG Y X, et al. FAPS: a fair, autonomous and privacy-preserving scheme for big data exchange based on oblivious transfer, Ether Cheque and smart contracts[J]. Information Sciences, 2021, 544: 469-484.
[12] JIANG P, GUO F C, LIANG K T, et al. Searchain: blockchain-based private keyword search in decentralized storage[J]. Future Generation Computer Systems, 2020, 107: 781-792.
[13] YANG H J, SHEN J, LU J Q, et al. A privacy-preserving data transmission scheme based on oblivious transfer and blockchain technology in the smart healthcare [J/OL]. Security and Communication Networks: 1-12[2021-12-04]. https://doi. org/10. 1155/2021/5781354.
[14] WANG X A, XHAFA F, MA J F, et al. Controlled secure social cloud data sharing based on a novel identity based proxy re-encryption plus scheme[J]. Journal of Parallel and Distributed Computing, 2019, 130: 153-165.
[15] 吴立强,韩益亮,杨晓元,等. 基于理想格的鲁棒门限代理重加密方案[J]. 电子学报, 2020, 48(9): 1786-1794.
WU L Q, HAN Y L, YANG X Y, et al. Robust threshold proxy re-encryption scheme from ideal lattices[J]. Acta Electronica Sinica, 2020, 48(9): 1786-1794. (in Chinese)
[16] LIU Z, SEO H, CASTIGLIONE A, et al. Memory-efficient implementation of elliptic curve cryptography for the Internet-of-things[J]. IEEE Transactions on Dependable and Secure Computing, 2019, 16(3): 521-529.
[17] 徐秋亮,李大兴. 椭圆曲线密码体制[J]. 计算机研究与发展, 1999, 36(11): 1281-1288.
XU Q L, LI D X. Elliptic curve cryptosystems[J]. Journal of Computer Research and Development, 1999, 36(11): 1281-1288. (in Chinese)
[18] BANSAL M, GUPTA S, MATHUR S. Comparison of ECC and RSA algorithm with DNA encoding for IoT security [C]//Proceedings of the 6th International Conference on Inventive Computation Technologies. Washington D. C., USA: IEEE Press, 2021: 1340-1343.
[19] RABIN M O. How to exchange secrets by oblivious transfer: TR-81[R]. Cambridge, USA: Aiken Computation Laboratory, Harvard University, 1981.
[20] 李顺东,杜润萌,杨颜璟,等. 安全多方多数据排序[J]. 计算机学报, 2020, 43(8): 1448-1462.
LI S D, DU R M, YANG Y J, et al. Secure multiparty multi-data ranking[J]. Chinese Journal of Computers, 2020, 43(8): 1448-1462. (in Chinese)
[21] CHOU T, ORLANDI C. The simplest protocol for oblivious transfer[M]. Berlin, Germany: Springer, 2015.
[22] 魏晓超,蒋瀚,赵川. 一个高效可完全模拟的n取1茫然传输协议[J]. 计算机研究与发展, 2016, 53(11): 2475-2481.
WEI X C, JIANG H, ZHAO C. An efficient 1-out-of-n oblivious transfer protocol with full simulation[J]. Journal of Computer Research and Development, 2016, 53(11): 2475-2481. (in Chinese)
[23] 赵圣楠,蒋瀚,魏晓超,等. 一个单服务器辅助的高效n取k茫然传输协议[J]. 计算机研究与发展, 2017, 54(10): 2215-2223.
ZHAO S N, JIANG H, WEI X C, et al. An efficient single server-aided k-out-of-n oblivious transfer protocol[J]. Journal of Computer Research and Development, 2017, 54(10): 2215-2223. (in Chinese)
[24] 徐彦蛟,李顺东,王道顺,等. 基于椭圆曲线公钥系统的不经意传输协议[J]. 计算机科学, 2013, 40(12): 186-191.
XU Y J, LI S D, WANG D S, et al. Oblivious transfer based on elliptic curve public key cryptosystems[J]. Computer Science, 2013, 40(12): 186-191. (in Chinese)
[25] GOYAL V, JAIN A, JIN Z Z, et al. Statistical zaps and new oblivious transfer protocols[C]//Proceedings of Annual International Conference on the Theory and Applications of Cryptographic Techniques. Berlin, Germany: Springer, 2020: 668-699.
[26] MOHAN A P, MOHAMED A A R, GLADSTON A. Merkle tree and blockchain-based cloud data auditing[J]. International Journal of Cloud Applications and Computing, 2020, 10(3): 54-66.
[27] DAVE J, DUTTA A, FARUKI P, et al. Secure proof of ownership using Merkle tree for deduplicated storage[J]. Automatic Control and Computer Sciences, 2020, 54(4): 358-370.
[28] 姜守旭,李建中. 一种P2P电子商务系统中基于声誉的信任机制[J]. 软件学报, 2007, 18(10): 2551-2563.
JIANG S X, LI J Z. A reputation-based trust mechanism for P2P E-commerce systems[J]. Journal of Software, 2007, 18(10): 2551-2563. (in Chinese)
[29] 温啸林,李长林,张馨艺,等. 基于DPoS共识机制的区块链社区演化的可视分析方法[J]. 计算机科学, 2022, 49(1): 328-335.
WEN X L, LI C L, ZHANG X Y, et al. Visual analysis method of blockchain community evolution based on DPoS consensus mechanism[J]. Computer Science, 2022, 49(1): 328-335. (in Chinese)
[30] GOLDREICH O. The fundamental of cryptography: basic applications[M]. London, UK: Cambridge University Press, 2004.

编辑 陆燕菲