基于区块链的信息共享及安全多方计算模型

王 童 马文平 罗 维
(西安电子科技大学通信工程学院 西安 710071)
(西安电子科技大学综合业务网国家重点实验室 西安 710071)

摘 要 在大数据背景下,数据信息隐私和可控性成为了关注点。现有的计算模式大多依赖于第三方机构,第三方的不可依赖性和对信息的掌控易导致信息的安全性无法得到保证,容易出现大量隐私问题。为解决此问题,文中结合区块链的特征和安全多方计算,提出了一种安全、高性能的共享及多方计算模型,使得用户能在自主控制数据的同时也能保证数据计算和共享的安全性。该方案首先以链上存储和链下存储相结合作为基础,在该存储环境下,利用代理重加密方式进行数据共享;然后使用改进的共识算法确保节点间的一致性,进而在MapReduce计算框架中使用改进的同态加密算法实现在无需解密隐私数据的情况下直接用密文进行数据处理和安全计算;最后对方案的正确性与安全性进行分析并进行实验仿真。分析结果及仿真结果表明,该模型在数据量较大时具有高性能的优点,且在运算效率方面有比较大的提升。

关键词 区块链,共识算法,MapReduce,同态加密,安全多方计算
中图法分类号 TP309.2 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2019.09.023

Information Sharing and Secure Multi-party Computing Model Based on Blockchain

WANG Tong MA Wen-ping LUO Wei
(School of Communication Engineering, Xidian University, Xi'an 710071, China)
(National Key Laboratory of Comprehensive Business Network, Xidian University, Xi'an 710071, China)

Abstract Under the background of big data, the control and privacy of data information have become a concern. However, existing computation models mostly rely on the third-party institution. Because the incompliance and the information control of the third party cause that information security cannot be guaranteed, more privacy problems appear. To solve this problem, this paper constructed an information sharing and secure multi-party computing model with high performance and security combining the blockchain with the secure multi-party computation, which enables users to control the data autonomously while ensuring the security of data information computing and sharing. This scheme firstly combines the on-chain storage with the off-chain storage. In this storage condition, proxy heavy encryption is used for data sharing and improved consensus algorithm is used to ensure the accuracy of nodes. Then, based on the MapReduce parallel computing framework, an improved homomorphic encryption algorithm was put forward for data processing and secure computing in cipher without decrypting the privacy data. Finally, the correctness and the security of the scheme were analyzed, and the experimental simulation was carried out. The analysis results and experimental results show that this scheme has high performance when dealing with big data and has a great improvement in operational efficiency.

Keywords Blockchain, Consensus algorithm, MapReduce, Homomorphic encryption, Secure multi-party computation

1 引言

随着当今社会的迅速发展,数字化和信息化的程度越来越高,数据隐私对人们的生活和工作产生了很大的影响,在医疗健康、金融等领域,数据隐私保护已成为重中之重。传统系统的中心化他信机制由于不基于完全可靠的第三方,数据隐私容易泄露;并且数据在第三方中也不能由用户自己控制,对数据的不可控性也会产生一系列的安全问题。

区块链技术具有去可信第三方的共信特性[1],区块链的数据对所有人公开,任何人都可以通过公开的接口查询数据,整个系统的信息高度透明,且数据拥有者(即用户)可以自己控制数据,而不是将数据交由不可信的第三方保管。本文针对个人信息在共享和计算中出现的安全问题和隐私问题[2],结合区块链的特征[3]和改进的同态加密算法,提出了基于区

到稿日期:2018-08-07 返修日期:2018-11-22 本文受国家自然科学基金(61373171),高等学校创新引智计划项目(B08038),国家重点研发计划重点专项(2017YFB0802400)资助。

王 童(1993-),女,硕士生,主要研究方向为信息安全和通信理论,E-mail:357146415@qq.com;马文平(1966-),男,教授,博士生导师,主要研究方向为密码学和信息安全,E-mail:357146415@qq.com(通信作者);罗 维(1987-),男,博士生,主要研究方向为密码学和云计算安全。

块链的共享及安全多方计算模型。

考虑到区块链不适于大规模数据分析计算的情况，提出了链上存储与链下存储相结合的方式来存储数据，将索引信息存储于链上，大规模数据信息则在链下进行存储，本文进行共享及计算的操作都是基于这种存储结构进行的。在信息计算共享的过程中，区块链去中心化的特点决定了可能会出现恶意节点以及因各方利益不一致导致的数据分歧等问题，这时需要利用共识算法来保障节点间的一致性。网络中常用的共识算法有工作量证明（POW）、股权证明（POS）、DPOS、RAFT 等。Castro 等[4] 于 1999 年提出了拜占庭容错算法，2016 年，黄步添等提出了一种基于区块链的拜占庭容错算法[5]，本文将这两个算法进行结合，设计了一种改进的拜占庭共识算法，解决了节点间相互信任的问题，为计算和共享提供了安全保证。

基于区块链构建的可行安全计算的平台[6]，考虑到对较大数据的处理会产生高性能计算复杂度高的问题，本文利用 MapReduce 框架来实现并行加密，并设计了一种改进的加密算法进行安全多方计算的优化，以保证数据的隐私性。1978 年，Rivest 等[7] 提出了同态加密的概念，在不知道明文的基础上可以对密文直接进行操作。2013 年，Plantard 等[8] 提出了基于理想格的全同态加密体制，在加密过程中，随着对密文操作的不断增加，噪声也不断加大。2010 年，文献[9]给出了 DGHV 方案，该方案支持整数上的加法运算和乘法运算。本文基于 DGHV 同态加密方案，利用 MapReduce 框架设计了一种改进的同态加密算法。

本文模型利用了区块链去中心化的特性，使用户能自主控制数据，使用区块链的共识算法保证了计算节点的一致性。在可信的环境下，基于 MapReduce 框架，在不泄露数据明文的情况下利用改进的同态加密算法对密文进行数据计算运行。即在区块链的基础上构建了一个可控安全的可行计算平台，使用户数据的隐私性和可控性得到保护，并在保护隐私的同时使数据得以共享和计算，最后进行大数据分析。

2　基于区块链的信息共享及安全多方计算模型框架

本节主要从以下几个方面实现数据信息的共享、大数据分析及隐私保护：数据的存储、查询者对数据的访问、不同机构对数据的安全多方计算等。图 1 给出了基于区块链的信息共享及安全多方计算模型框架。

数据层中存放的是用户的原始隐私数据，由用户自己保管，用户不仅可以上传数据而且具有对数据的控制权。

在存储层中，考虑到数据在区块链上存储，区块同步时也会将大量数据在区块中进行同步，这会占用空间，从而导致资源浪费。因此，本方案将链上索引表信息与链下数据库相结合进行存储，一方面能释放区块链上的大量空间，另一方面还能提高信息共享的效率。

链上的索引表存储索引信息（索引类别及加密文件的地址），并形成索引区块存放在区块链上。索引表将查询者查询的信息类别与存储地址值相对应，在链上占用少部分内存。链下的数据库存储数据拥有者上传的加密数据文件，保证数据的安全性。

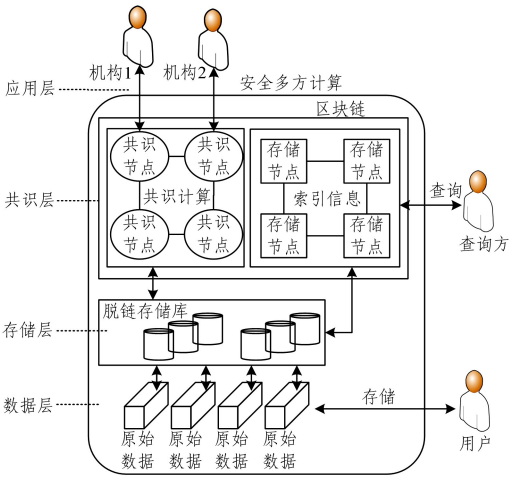

图 1　基于区块链的信息共享及多方计算的隐私保护框架

Fig. 1　Privacy protection framework of sharing and computing information based on blockchain

在信息共享时，如果查询者想要查询相关数据信息，可以向区块链发出信息查询请求，经授权后可以通过索引表中的地址值在链下数据库中查找加密文件，利用用户私钥和查询者公钥生成的重加密密钥[10] 对加密文件进行二次加密，并将加密结果返回至查询者，用查询者的私钥进行解密获得信息。

在应用层中，当机构需要对数据进行分析统计，以进行更好的研究时，将采用安全多方计算的方式。计算请求方生成计算请求后，请求方将自己的证书签名和加密公钥提交到区块链上，利用共识层的共识节点进行共识来确保节点间的相互信任，之后根据请求的内容查询原始数据，并将不同机构的数据合并在一起进行计算，利用请求方的公钥将计算的结果加密返回。在此过程中，每一个机构均不知道其他请求方的信息，请求方收到返回的结果后，利用自己的私钥进行解密得到最终的结果。这种方式在信息隐私保护的前提下实现了数据联合共享计算的功能[11]。

3　模型的详细设计

该模型利用区块链技术和安全多方计算技术共同进行数据信息的隐私保护。该模型分为三部分：存储部分、共识部分及计算部分。存储过程分为链上的索引存储和链下的大部分数据存储两个部分[12]，敏感的信息及数据量较大的信息在链下进行存储，链上存储用户及文档的索引信息及数据量较小的信息。在进行数据计算的过程中，在区块链上先利用共识机制解决区块链节点间相互信任的问题，在相互信任的基础上再利用安全多方计算技术使各节点共同计算出需要得到的函数值，在此过程中不会泄露任何隐私信息。

3.1　区块链的存储模型

考虑到数据量较大的信息在区块链上存储会造成内存的过度占用，尤其当区块进行同步时信息会在各个节点中进行同步，这在一定程度上会导致资源的浪费和负担太重的问题，因此敏感及数据量较大的数据在链下以密文形式进行存储。不论是用户本身还是需要使用数据的机构，都不能随意更改数据，由此保证了数据的安全性。

图 2 给出了链上索引区块的数据结构。索引区块由区块

头和区块体构成。区块头里存储着区块的头信息，包含上一个区块的哈希值、本区块体的哈希值以及时间戳等信息。区块体中存储着查询者需要查询的索引信息，并通过区块之间的连接进行区块的同步。

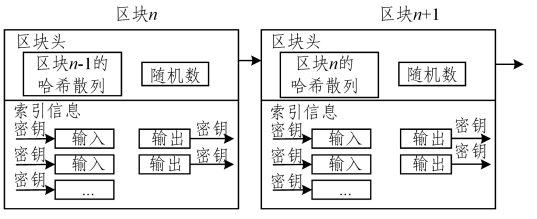

图 2 链上索引区块的数据结构

Fig. 2 Data structure of index block on chain

在存储阶段：用户对要上传的文件信息进行分类，如商业信息、医疗信息等，并对外公布分类，然后将对应分类中的数据用公钥进行加密并存储到数据库中形成加密文件，之后将存储的地址值返回至链上并存储在索引表中。

在数据共享阶段：利用代理重加密机制[13]来保证在不将用户私钥直接暴露给查询者的前提下对数据进行共享。用户会产生对应于用户自己到查询者的代理重加密密钥，并将重加密密钥发送给链上的任意节点。代理重加密节点根据查询者需要查询的密文和重加密密钥完成重加密操作并将加密数据传送给查询者，查询者利用自己的私钥对数据进行解密以获取数据从而完成共享操作。具体过程如图 3 及算法 1 所示。

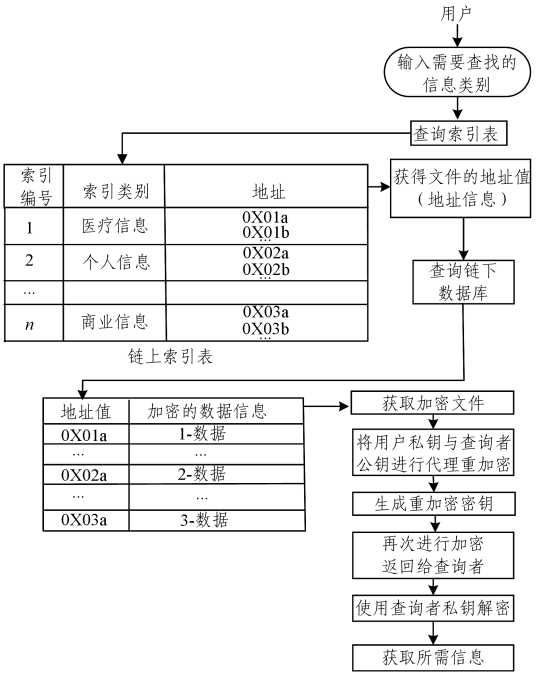

图 3 数据存储及共享流程图

Fig. 3 Flowchart of data storage and sharing

算法 1 数据信息存储共享

Procedure Storing(数据信息 m)

Input：数据信息 m，查询者的公钥 pk 及私钥 sk，使用者请求读取的操作 rw

1. begin
2. if rw=0（rw=0 为存入数据，rw=1 为读出数据）then
3. 数据拥有者将上传数据进行分类(t_1，t_2，…)
4. return 数据类别 then
5. 将不同类别的数据 m 用公钥进行加密
6. 存储至链下数据库
7. return 不同类别信息存储的地址值 to 链上 then
8. 创建索引表
9. return 索引编号
10. if rw=1 then
11. 查询者在链上查询所需的数据类别
12. 在链下数据库查询地址值对应的加密信息 c=Enc(m)
13. 用户利用自己的私钥和查询者提供的公钥进行代理重加密以生成重加密密钥 k
14. 重加密节点利用重加密密钥对密文数据进行加密生成 $Enc_k(c)$
15. return 加密信息 $Enc_k(c)$ to 查询者
16. 查询者利用自己的私钥 sk 进行解密
17. return 明文信息
18. end

3.2 共识算法设计

在区块链系统中，消息可能会出现丢失、损坏、传输错误等问题，此外去中心化的特点决定了系统中任何一个参与者都不能被信任，可能会产生恶意节点的出现和各方利益不一致的问题。因此，为了防止此类错误的发生，需要利用一个共识机制来保证节点间的一致性，并确保每个节点都有一个公认的全局账本。

传统的共识机制只是针对某些特定问题的容错方法，并不能完全解决区块链系统的容错问题。基于对不同共识算法的比较和分析，本文设计了一种改进的共识算法，在此基础上实现云计算环境下的去中心化的安全多方计算，即改进的拜占庭容错算法，并将其应用于区块链系统中。本文的共识算法由 n 个共识节点组成，提供了 $f=(n-1)/3$ 的容错能力，这种容错能力同时具有安全性和可用性。在改进的 PBFT 中，节点同步过程采用向其他节点索要区块并校验的方式完成同步。视图切换协议在结合区块生成协议的基础上采用超时机制进行试图切换，一定程度上减少了算力开销和数据传输量。图 4 为共识算法的流程图。

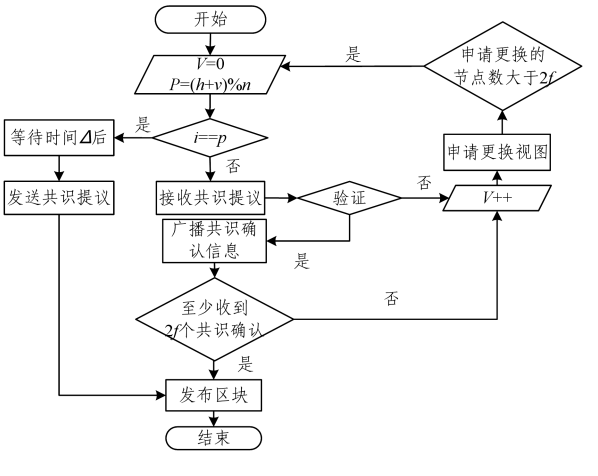

图 4 共识算法的流程图

Fig. 4 Flowchart of consensus algorithm

3.2.1 共识流程

(1)系统中交易的发起者发起一笔交易时，利用私钥签名进行交易后，向全网进行广播。

(2)当节点收到一笔交易后，判断该节点是否为主节点。若不是主节点，则将交易内容转发给下一个节点。若为主节点，则需要验证交易的合法性，若合法，则将这个交易写到区块体的交易字段中。

(3)主节点经过时间 t 发送共识提案$\langle PrepareRequest,h,v,P,block,\langle block\rangle_{\sigma p}\rangle$。

(4)其他节点在收到主节点发送过来的共识提案后，检查提案为真后，向除自己之外的从节点及其主节点发送共识确认消息$\langle PrepareResponse,h,v,i,block,\langle block\rangle_{\sigma i}\rangle$。

(5)若其他节点对共识提案校验后发现不为真时，则主节点就可能被其他节点怀疑，从而进行试图更换的广播。

(6)网络中任意节点在收到 $2f$ 个相同的确认消息后，即认为共识达成，可以发布来自主节点的区块 $block$。

(7)其余节点在收到 $block$ 后，认为该轮共识任务完成，将自己存储在区块里的交易删除，并开始下一轮共识。

表 1　共识算法中的变量符号说明

Table 1　Variable symbol description in consensus algorithm

符号	含义
n	全网中参与共识的节点总数
R	共识节点的集合
f	网络中允许错误节点的最大数
v	视图的编号
h	区块的高度
p	议长编号

3.2.2　视图更换流程

当节点 i 在经过 $2^{v_k+1}\cdot t$ 时间后仍未达成共识，或接收到含有非法交易的提案时，进入试图更换流程：

(1)令 $k=1$，$v_k=v+k$；

(2)节点 i 发出视图更换请求$[ChangeView,h,v,i,v_k]$；

(3)任意节点收到至少 $2f$ 个来自不同 i 的相同 v_k 后，视图更换达成，令 $v=v_k$，开始共识；

(4)如果在 $2^{v_k+1}\cdot t$ 时间间隔后视图仍未达成，则 k 递增并返回到步骤(2)。

3.3　基于 MapReduce 的安全多方计算模型

3.3.1　MapReduce 编程模型

在确保节点间一致性的基础上，利用 MapReduce 进行同态加密运算。MapReduce 计算框架是一种运行在集群节点上的能处理海量数据的并行编程模型和框架[14]。当需要处理的数据量很大时，利用 MapReduce 可以提高数据计算的效率。图 5 给出了 MapReduce 的框架流程，MapReduce 可以分为 Map 和 Reduce 两个阶段，主节点对节点进行 Map 和 Reduce 的任务分配。

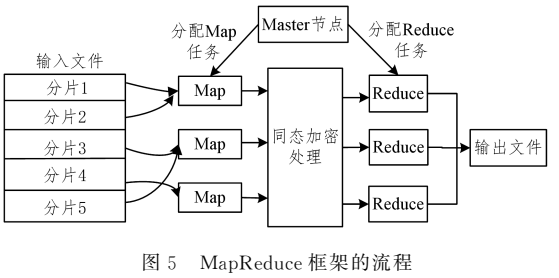

图 5　MapReduce 框架的流程

Fig. 5　Process of MapReduce framework

图 6 是区块链上的节点参与安全多方计算的流程图。在加密过程中，利用 MapReduce 框架判断节点是否为主节点，若是主节点则由主节点分配 Map 任务和 Reduce 任务，若不是主节点则判断该节点是 Map 节点还是 Reduce 节点。如果是 Map 节点，则利用改进的同态加密算法对分片后的数据进行数据的并行加密；如果是 Reduce 节点，则对加密后的信息进行汇总处理从而得到输出数据。最后将加密后的密文存入链下，由用户利用公钥自己控制数据。只有经过用户同意授权的使用者才可以获取真实的数据。

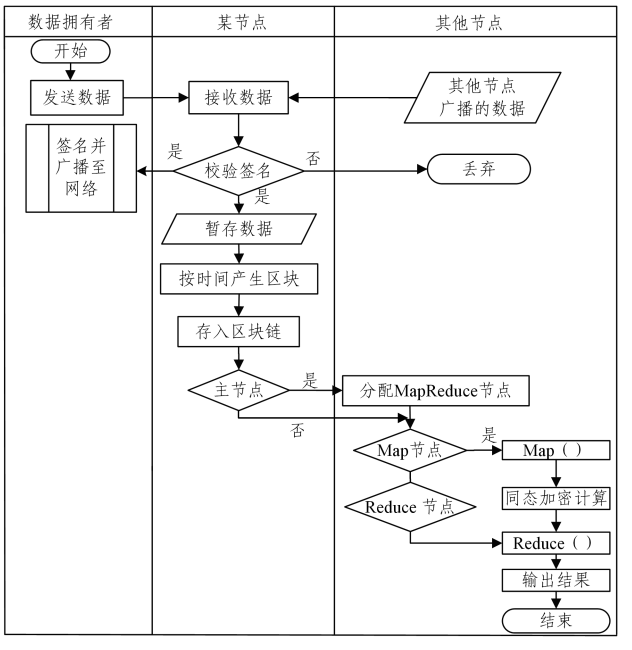

图 6　区块链上的节点参与安全计算的流程图

Fig. 6　Flowchart of nodes on blockchain participating in security computation

3.3.2　改进的同态加密算法

在传统的公钥加密方案中，一般使用幂运算或者模运算来进行加密，虽然在安全性上都可以得到保证，但是这些运算都是较复杂的运算。文献[9]提出的 DGHV 算法是在整数的范围内实现了加法同态和乘法同态。但在提出过的基于整数的同态加密算法中，有的算法只能加密一个比特的数据，有的加密运算需要大量的空间来存储所需要的公钥，加密时间长，运算效率低。

在本文改进的同态加密算法中，加密明文的比特数为 k，且公钥中分量的形式采用三次方形式，相对来说得到了相对最小的密文长度和公钥尺寸，在数据量较大时，节省了系统加密的时间，提高了算法的运算效率。

本文的加密方案由 4 个部分组成：密钥生成算法 KeyGen()、加密算法Encrypt()、解密算法()和评估算法 Evaluate()。表 2 列出了同态加密过程中需要用到的变量符号的说明。

表 2　同态加密中的变量符号说明

Table 2　Variable symbol description in homomorphic encryption

符号	含义
r	公钥中整数的比特长度
η	私钥的比特长度
ρ	噪声的比特长度
τ	公钥中整数的个数
ρ'	第二个噪声参数

首先，按照文献[9]中 Gentry 的构造思路，构造一个 Somewhat 同态方案。但本文构造的是一个整数 $x'_{i,j,k}$，$x'_{i,j,k}=x_{i,0}\cdot x_{j,1}\cdot x_{k,2} \bmod x_0$，且 i,j,k 满足 $1\leqslant i,j,k\leqslant\beta$(其中 β 是一个新的参数)。这样为了产生 $\tau=\beta^3$ 个 $x'_{i,j,k}$ 用于加密，只须在公钥中存储 3β 个 $x_{i,b}$，公钥尺寸也就从 τ 下降为 $\sqrt[3]{\tau}$；并且该算法的模 2 运算改成了模 2^k 运算，改进后的算法可以一次加密 $k$bit。参数 $\rho=\lambda,\rho'=2\lambda$，其中 λ 为安全参数

KeyGen(1^λ)：随机产生一个 ηbit 素数 p，$p\in[2^{\eta-1},2^\eta)$，令 $x_0=q_0\cdot p$，q_0 是 $[0,2^\gamma/p)$ 中的奇整数。产生整数 $x_{i,b}=p\cdot q_{i,b}+2^k\cdot r_{i,b}$ 且 $1\leqslant i\leqslant\beta,0\leqslant b\leqslant 2$，其中 $q_{i,b}$ 是 $c=m+2^k r+2^k\sum\limits_{1\leqslant i,j,k\leqslant\beta} b_{i,j,k}\cdot x_{i,0}\cdot x_{j,1}\cdot x_{k,2}+kx_0$ 的随机整数，$r_{i,b}$ 是 $(-2^\rho,2^\rho)$ 中的整数。令 $sk=p$，$pk=(x_0,x_{1,0},x_{1,1},x_{1,2},\cdots,x_{\beta,0}x_{\beta,1}x_{\beta,2})$。

Encrypt($pk,m\in\{0,1\}^k$)：随机产生一个尺寸为 $\tau=\beta^3$ 的向量 $b=(b_{i,j,k})$。向量系数在 $[0,2^\alpha)$ 产生一个随机整数 $r\in(-2^{\rho'},2^{\rho'})$。输出密文为：$c=m+2^k r+2^k\sum\limits_{1\leqslant i,j,k\leqslant\beta} b_{i,j,k}\cdot x_{i,0}\cdot x_{j,1}\cdot x_{k,2} \bmod x_0$。

Evaluate($pk,C,c_1,\cdots,c_t$)：给定一个 t 输入的二元门电路 C 和 t 个密文 C_i，将 C 中的加法和乘法门电路用于密文操作，所有操作在整数上执行，返回结果为整数。

Decrypt(sk,c)：输出 $m\leftarrow(c \bmod p) \bmod 2^k$。

(1)算法正确性证明

存在一个整数 k，使得密文 $c=m+2^k r+2^k\sum\limits_{1\leqslant i,j,k\leqslant\beta} b_{i,j,k}\cdot x_{i,0}\cdot x_{j,1}\cdot x_{k,2}+kx_0$，$|k|\leqslant\tau$。因为 $x_{i,b}=p\cdot q_{i,b}+2^k\cdot r_{i,b}$，$x_0=q_0\cdot p$，所以有 $c=m+2^k r+2^k\sum\limits_{1\leqslant i,j,k\leqslant\beta} b_{i,j,k}\cdot r_{i,0}\cdot r_{j,1}\cdot r_{k,2} \bmod p$

由于 $\rho'\geqslant 3\cdot\rho+\alpha$，因此有 $|c \bmod p|\leqslant 2^{\rho'+k}+2^k\cdot\tau\cdot 2^{3\rho+\alpha}\leqslant\tau\cdot 2^{\rho'+k+1}$。

令 $C\in C_\varepsilon$ 是一个有 t 个输入的可操作电路，C' 是相应的在整数上操作的电路。已知 $c_i\leftarrow Encrypt(pk,m_i)$，可得 $c \bmod p=C'(c_1,\cdots,c_t) \bmod p=C'(c_1 \bmod p,\cdots,c_t \bmod p) \bmod p$。由可操作电路的定义可得 $|C'(c_1 \bmod p,\cdots,c_t \bmod p)|\leqslant(2^{\eta-4})^{k-1}\leqslant(p/8)^{k-1}$，

因此 $C'(c_1 \bmod p,\cdots,c_t \bmod p) \bmod p=C'(c_1 \bmod p,\cdots,c_t \bmod p)c \bmod p=C'(c_1 \bmod p,\cdots,c_t \bmod p)$，

$[c \bmod p]_{2^k}=[C'(c_1 \bmod p,\cdots,c_t \bmod p)]_{2^k}=C'([c_1 \bmod p]_{2^k},\cdots,[c_t \bmod p]_{2^k})$，$[c \bmod p]_{2^k}=C(m_1,\cdots,m_t)$，算法正确性得证。

(2)输出同态性分析

加法同态：

设有两个消息 m_0 和 m_1，对他们分别加密后，可以得到：

$$c_0=m_0+2^k r_0+2^k\sum_{1\leqslant i,j,k\leqslant\sqrt[3]{\tau}} b_{i,j,k}\cdot x_{i,0}\cdot x_{j,1}\cdot x_{k,2} \bmod x_0$$

$$c_1=m_1+2^k r_1+2^k\sum_{1\leqslant i,j,k\leqslant\sqrt[3]{\tau}} b_{i,j,k}\cdot x_{i,0}\cdot x_{j,1}\cdot x_{k,2} \bmod x_1$$

必然存在两个整数 k_0 和 k_1，使得：

$$c_0=m_0+2^k r_0+2^k\sum_{1\leqslant i,j,k\leqslant\sqrt[3]{\tau}} b_{i,j,k}x_{i,0}x_{j,1}x_{k,2}+k_0q_0p$$

$$c_1=m_1+2^k r_1+2^k\sum_{1\leqslant i,j,k\leqslant\sqrt[3]{\tau}} b'_{i,j,k}x'_{i,0}x'_{j,1}x'_{k,2}+k_1q_1p$$

从而可以得到：

$$c_0+c_1=m_0+m_1+2^k(r_0+r_1+\sum_{1\leqslant i,j,k\leqslant\sqrt[3]{\tau}} b_{i,j,k}x_{i,0}x_{j,1}x_{k,2}+\sum_{1\leqslant i,j,k\leqslant\sqrt[3]{\tau}} b_{i,j,k}{}'x_{i,0}{}'x_{j,1}{}'x_{k,2}{}')+p(k_0q_0+k_1q_0)$$

对密文进行解密，可以得到：

$$\begin{aligned}m&=((c_0+c_1) \bmod p) \bmod 2^k\\&=[2^k(r_0+r_1+\sum_{1\leqslant i,j,k\leqslant\sqrt[3]{\tau}} b_{i,j,k}x_{i,0}x_{j,1}x_{k,2}+\sum_{1\leqslant i,j,k\leqslant\sqrt[3]{\tau}} b_{i,j,k}{}'x_{i,0}{}'x_{j,1}{}'x_{k,2}{}')+p(k_0q_0+k_1q_0)] \bmod p \bmod 2^k\\&=m_0+m_1\end{aligned}$$

因此，该算法满足加法同态。

乘法同态：

根据加法同态的证明，可以将密文写成如下形式：

$$\begin{aligned}c_0&=m_0+2^k r_0+2^k\sum_{1\leqslant i,j,k\leqslant\sqrt[3]{\tau}} b_{i,j,k}x_{i,0}x_{j,1}x_{k,2}+p(k_0q_0)\\&=m_0+2^kA_0+PB_0\\c_1&=m_1+2^k r_1+2^k\sum_{1\leqslant i,j,k\leqslant\sqrt[3]{\tau}} b_{i,j,k}{}'x_{i,0}{}'x_{j,1}{}'x_{k,2}{}'+p(k_1q_0)\\&=m_1+2^kA_1+PB_1\end{aligned}$$

$$c_0\times c_1=(m_0\times m_1)+2^kA+PB$$

$$m=((c_0\times c_1) \bmod p) \bmod 2^k=m_0\times m_1$$

因此该算法满足乘法同态。

(3)算法安全性证明

本方案的安全性归约于近似最大公因子问题，也就是针对该方案的攻击都可以转换为最大公因子问题，这一点与 DGHV 方案相似，其安全级别达到了 IND-CPA 安全，在文献[9]中已经得到论证。目前为止，最大公因子问题是不能破解的，因此该方案符合安全性。

由此可见，改进的同态加密算法同时满足同态性、正确性、安全性。

表 3 几种同态方案的比较

Table 3 Comparison of several homomorphic schemes

算法	加密明文比特/bit	公钥中分量的形式	公钥尺寸	私钥尺寸	数题难题
DGHV	1	线性	$\tilde{o}(\lambda^{10})$	$\tilde{o}(\lambda^2)$	近似最大公因子问题
文献[15]	k	线性	$\tilde{o}(\lambda^7)$	$\tilde{o}(\lambda^2)$	近似最大公因子问题
文献[16]	1	二次	$\tilde{o}(\lambda^7)$	$\tilde{o}(\lambda^2)$	近似最大公因子问题
文献[17]	k	线性	$\tilde{o}(\lambda^5)$	$\tilde{o}(\lambda^2)$	部分近似最大公因子问题
本文算法	k	三次	$\tilde{o}(\lambda^4)$	$\tilde{o}(\lambda^2)$	近似最大公因子问题

由表 3 可以看出，DGHV 算法的密钥开销最大，因为它是对单个比特进行加密的，而且公钥中分量形式为线性，所以加密得到的密文会很大。而文献[15-17]在一定程度上改变了加密明文的比特数或者改变了公钥中分量的形式，从而使密文长度变小，公钥尺寸相对变小。本文提出的改进的同态加密算法，加密明文比特数为 k，且公钥中每个分量的形式采用三次方形式，相对于以上几种加密算法，得到了相对最小的密文长度和公钥尺寸，在数据量特别大时，更适合大数据的处理。

(4)实验分析

本实验在 DGHV 算法的基础上进行了不同程度的改进,然后对 3 种算法进行并行加密得出运算时间随 Map 节点个数的变化的变化情况。实验中对 128 M 的文件信息进行加密,利用改进的算法一次加密 8 个字节,将安全参数设为 2。

由图 7 可以得出两条结论:1)对于每一种算法,运算所耗费的时间随着 Map 节点个数的增加而减少,且在 Map 节点增加至一定值时,会趋于平稳,因为随着 Map 节点个数的增加,信息会被分成更多数据块一起参与运算,提高了运算的速率;2)在 DGHV 算法的基础上,当加密明文比特数为 k,公钥中分量采用三次方形式时,所耗费的时间最少,这是因为增加了加密明文的比特数或者增加了公钥中分量的形式,使得密文长度变小,从而缩短了运算时间。

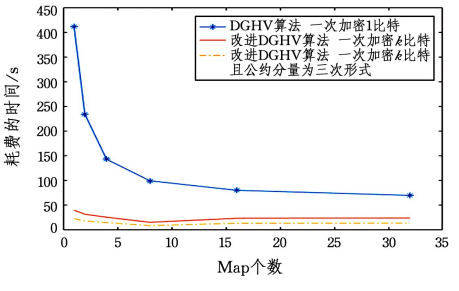

图 7　3 种算法的运算时间与 Map 节点个数的关系图

Fig. 7　Relation between computation time of three algorithms and number of Map nodes

4　安全性分析

本节针对普通的安全多方计算与本文基于区块链的安全多方计算在交易验证、安全存储、用户控制等方面进行了综合比较。

4.1　交易验证

共识机制是区块链的一个重要特征。传统的安全多方计算可能存在参与方共谋篡改数据的安全隐患。以区块链作为安全多方计算的基础,在计算之前,节点同步过程采用向其他节点索要区块并校验的方式完成同步,保证节点间的一致性,进而才能进行同态加密计算,以保证交易的安全性。

4.2　防篡改

区块链是公开的账本,区块链中的所有区块都按时间的先后顺序排列,每个区块都包含了上一个区块的信息,要修改某个数据至少需要 51%以上的算力,因此数据信息一旦写入到区块链上,就无法被篡改。

4.3　安全存储

由于区块链是公开的账本,公共信息在链上存储无法篡改,因此存储的安全性得到了保证。数据文件放在链下的存储库中。在存储之后,构建索引表并将文件的存储地址与索引编号作为索引信息并形成索引区块存储在链上,链上的索引表对应的地址信息用于在链下数据库中对文件进行查找。

4.4　用户控制

在本文中,用户不依赖于可信的第三方,结合安全多方计算,用户可以自主控制自己的数据信息,具有对数据的所有权,用户可以同意数据使用者对自己数据进行访问,也可以撤销其访问权。只有用户同意请求者的访问请求,查询者和使用者才可以进行数据查找及计算。

4.5　方案比较

本文提出了基于区块链的共享及安全多方计算模型。该模型基于区块链去中心化的特性,在进

行数据共享的同时利用安全多方计算模型进行隐私保护。表 4 列出了现有的计算模型与本文提出的模型在安全性方面的比较,表 5 列出了面临的问题及利用模型解决的方法。

表 4　两种安全多方计算的安全性分析

Table 4　Security analysis between two secure multi-party computations

	依赖可信第三方	交易验证	安全存储	用户控制权	隐私保护
现有的安全多方计算	√	×	×	×	√
基于区块链的安全多方计算	×	√	√	√	√

表 5　面临的问题及利用模型解决的方法

Table 5　Problems and solutions

类型	问题	模型应对方法
隐私保护	1. 参与方中会有恶意节点的出现 2. 在共享数据时对数据信息的保护	1. 区块链共识机制确保节点间数据一致 2. 采用改进的同态加密算法进行安全多方计算来加密数据,在密文的基础上进行计算
用户控制权	用户无法控制自己的数据信息	用户具有控制权,可以同意或撤销请求者访问数据的权限
数据篡改滥用	1. 数据被滥用,追责困难 2. 数据容易被篡改	1. 区块链是维护数据记录的分布式账本,通过链式结构可追溯到上一区块中的记录 2. 区块链中每个区块都包含了上一个区块的信息,与之前信息相关联,难以篡改
数据计算效率	当信息数据较大时,计算效率缓慢	MapReduce 计算框架的并行处理结构适应大数据处理
数据存储	1. 大规模数据存储时,区块链采用冗余方式不适合存储 2. 数据容易丢失	1. 采用区块链链上和链下结合的数据存储方式 2. 分布式的存储节点可以用于实时共享数据

结束语　数据隐私泄露的问题对当今社会造成了非常大的影响,单纯地使用安全多方计算在一定程度上不会泄露用户的隐私,但是用户不能自主控制数据,第三方机构很有可能因自身利益获取数据。区块链技术结合安全多方计算,可以保证数据共享的同时隐私问题得到进一步的改善。

(1)利用区块链去中心化的特点,公共信息保存在区块链上且无法被篡改,共识机制确保节点的一致性和安全性,区块链分布式的结构使得数据分散在不同的节点上,不易受到攻击,单点威胁不会影响到整个网络。

(2)安全多方计算在理论上被证明是安全的,确保在用户共享数据的过程中不泄露原始数据,改进的同态加密算法也为用户的数据隐私的安全性提供了保障。

在将区块链应用于安全多方计算的同时,还有许多需要完善的地方,如在金融医疗等各个应用场景中,需要考虑实际的业务流程对方案进行详细设计和细节描述。还要建立一种

数据标准来保证数据的兼容性,从而保证在使用者和数据拥有者之间有正常的数据流通。

参 考 文 献

[1] ZHU L H,GAO F,SHEN M,et al. Survey on Privacy Preserving Techniques for Blockchain Techniques[J]. Computer engineering and Application,2017,54(10):2170-2186. (in Chinese)
祝烈煌,高峰,沈蒙,等. 区块链隐私保护研究综述[J]. 计算机研究与发展,2017,54(10):2170-2186.

[2] DORRI A,STEGER M,KANHERE S S,et al. BlockChain: A Distributed Solution to Automotive Security and Privacy[J]. IEEE Communications Magazine,2017,55(12):119-125.

[3] LI X,JIANG P,CHEN T,et al. A Survey on the security of blockchain systems[J]. Future Generation Computer Systems,2017:1-13. doi:10.1016.

[4] CASTRO M,LISKOV B. Practical Byzantine fault tolerance[C]// Symposium on operating Systems Design & Implementation. ACM,1999,173-186.

[5] 黄步添,王云霄,王从礼,等. 一种应用于区块链的拜占庭容错共识方法:中国,CN106445711A[P]. 2017-02-22.

[6] ZYSKIND G,NATHAN O,ALE X,et al. Decentralizing Privacy:Using Blockchain to Protect Personal Data[C]// IEEE Security and Privacy Workshops. IEEE Computer Society,2015:180-184.

[7] RIVEST R L,ADLEM A L,DERTOUZOS M L. On Data Banks and Privacy Homomorphism [C]// Foundations of Secure Computation. New York:Academic Press,1978:169-179.

[8] PLANTARD T,SUSILO W,ZHANG Z. Fully Homomorphic Encryption Using Hidden Ideal Lattice[J]. IEEE Transactions on Information Forensics and Security,2013,8(12):2127-2137.

[9] DIJK M V,GENTRY C,HALEVI S,et al. Fully Homomorphic Encryption over the Integers[J]. Lecture Notes in Computer Science,2010,2009(4):24-43.

[10] TAN Z L,ZHANG W. Multiparty Cloud Computation and Homomorphic Proxy Re-encrypt Scheme[J]. Journal of Chinese Computer Systems,2015,36(8):1739-1742.

[11] FU D,FANG L. Blockchain-based trusted computing in social network[C]// IEEE International Conference on Computer and Communications. IEEE,2017:19-22.

[12] DO H G,NG W K. Blockchain-Based System for Secure Data Storage with Private Keyword Search[C]// Services. IEEE,2017:90-93.

[13] ZHENG Z H,ZHANG M Q,WANG X A. Identity based proxy re-encryption scheme for secure cloud data sharing[J]. Computer Engineering and Application,2016,33(11):3450-3454. (in Chinese)
郑志恒,张敏情,王绪安. 一种适合云数据共享的身份代理重加密方案[J]. 计算机应用研究,2016,33(11):3450-3454.

[14] DITTRICH J,QUIANÉ-RUIZ J A. Efficient big data processing in Hadoop MapReduce[J]. Proceedings of the Vldb Endowment,2012,5(12):2014-2015.

[15] XIE X S. A full homomorphic encryption scheme that is valid for a class of integers[D]. Jinan: Shandong university,2014. (in Chinese)
谢学说. 一类整数上有效的全同态加密方案[D]. 济南:山东大学,2014.

[16] TIBOUCHI M,MANDAL A. Fully Homomorphic Encryption over the Integers with Shorter Public Keys[C]// Conference on Advances in Cryptology. Springer-verlag,2011:487-504.

[17] TANG D H,ZHU S X,CAO Y F,et al. A full homomorphic encryption scheme on a faster integer[J]. Computer Engineering and Application,2012,48(28):117-122. (in Chinese)
汤殿华,祝世雄,曹云飞,等. 一个较快速的整数上的全同态加密方案[J]. 计算机应用与研究,2012,48(28):117-122.