

硕 士 研 究 生 读 书 报 告



题目 区块链隐私保护研究与实践综述

作者姓名 赵正

作者学号 22151099

指导教师 程学林

学科专业 软件工程

所在学院 软件学院

提交日期 二零二壹年十二月

Survey of Research and Practices on Blockchain Privacy Protection

A Dissertation Submitted to Zhejiang University in partial fulfillment of the requirements for the degree of Master of Engineering

Major Subject: Software Engineering

Advisor: Cheng Xuelin

By

Zhao Zheng

Zhejiang University, P.R. China

2021

**摘要**

基于区块链的分布式账本集成了非对称加密体系、P2P 网络、共识算法、智能合约等多种技术,保证事务记录的一致性和不可篡改性.但是,区块链技术中的账本共享机制也带来了隐私威胁,用户身份、账户地址、交易内容等信息的隐私保护成为研究的关注点.讨论了区块链系统中的隐私威胁;着重分析了地址混淆、信息隐藏、通道隔离等 3 类隐私保护机制,详细介绍各类机制的原理、模型、特征及实现技术;最后探讨了实际应用中,区块链隐私保护技术在系统性能和可扩展性方面的挑战和发展方向.

**关键词**：区块链，隐私保护，通道隔离

**Abstract**

Blockchain-based distributed ledger aims to provide consistent and tamper-resistant transaction records by integrating various security technologies such as asymmetric cryptosystem, P2P network, consensus algorithm, and smart contract. However, as each node in the blockchain system shares a copy of the public ledger, such data sharing mechanism also introduces vulnerabilities that hackers could exploit to attack private information. Privacy protection of blockchain systems thus gains wide attentions from researchers. Various techniques have been proposed to protect users’ identity, address, and transaction information from security threats. This study investigates blockchain privacy threats. It made a comprehensive survey of state-of-the-art privacy protection technologies which are categorized into three mechanisms including address confusion, information hiding, and channel isolation. The paper introduces the principles, models, and various implementations of each mechanism. It finally discusses the challenges of performance and scalability in practice and future technology advancement directions.

**Keywords：**Blockchain，Privacy protection，Channel isolation

# **区块链隐私及威胁**

传统的区块链系统中通常采用假名机制和广播机制保护用户隐私[1],其中,

• 假名机制指用户可以独立生成任意数量的区块链地址,不需要通过注册或者认证机制.同一用户生成的不同地址可以单独使用,彼此间不存在任何关联关系.因此,仅通过区块链地址无法关联到用户的真实身份,该机制能够隔离用户在区块链上不同操作的记录.

• 广播机制指区块链系统通过P2P网络传输数据,网络中采用洪水广播协议传播消息,接受节点无法判断消息来源是消息的直接发起者还是转发者,从而保护消息真实发起者的身份. 假名机制和广播机制能够在一定程度上保护区块链用户的隐私安全,但在实际应用中,用户隐私仍面临各类威胁,主要存在于记录数据的分布式账本和区块链去中心化网络中各节点的相关信息.为了保证去中心化系统的正确性和安全性,区块链系统中的所有节点共同维护一致的分布式账本,记录区块链系统中的所有历史数据,用于验证用户提交的新事务的合法性.为了所有节点都能验证账本的正确性,账本中所有数据保持公开,因此账本数据能够被攻击者轻易获取,攻击者通过分析公开账本中的记录严重威胁用户隐私.此外,区块链系统采用去中心化网路进行通信,在非许可链系统中,节点加入网络不需要任何身份认证,这在增强了扩展性的同时也导致攻击者可以自由部署节点加入网络,监听网络中各节点隐私信息以及网络中通信信息.

# **地址混淆机制**

区块链技术较为广泛应用于密码货币领域,由于区块链账本记录了历史上所有的交易记录,通常对网络中所有节点公开可见,因此,攻击者可以通过分析账本进行攻击.假设同一交易的所有输入地址属于同一用户.为了抵抗账本分析技术,研究者针对该技术所基于的假设,提出交换资产、混淆地址的防御机制,即地址混淆机制.不同的用户通过交易相互交换资产,这会将不同用户的账户地址误认为属于同一用户,达到混淆用户地址、保护各用户隐私的效果.由于地址混淆机制通过交换资产的方式进行,因而通常称为混币机制,用于交换资产的交易称为混币交易.

地址混淆机制有多种不同的实现形式,根据具体操作者的不同,分为中心化混币和去中心化混币两类技术:中心化混币技术需要中心化混币服务提供商参与,帮助混币用户进行混币操作;去中心化混币技术由所有参与混币的用户按照协议自发进行混币交易.

## 中心化混币

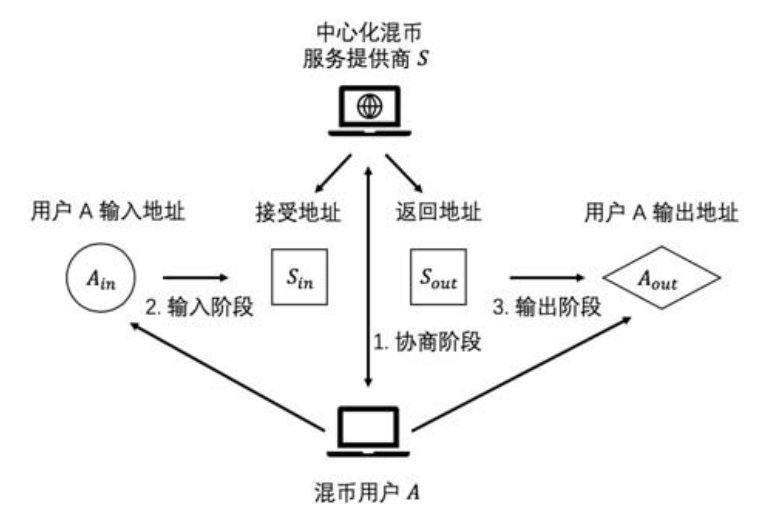
中心化混币服务提供商帮助希望进行混币交易的用户找到同伴,构造混币交易,并从中收取一定额度的手续费.中心化混币技术中,混币服务提供商作为中介角色分别与各用户进行交易,接收到用户的资产后,进行随机混淆,然后返回给其他用户.通过将不同用户的资产互相交换,达到混淆不同用户地址的效果,因而分析攻击只能将所有参与混币服务的地址聚类到一起,难以分辨出属于单一用户的账户地址.中心化混币协议的基本模型如下图所示,具体的协议流程主要分为协商、输入、输出及结束这 4 个阶段.

1. 协商阶段:希望参与混币的用户与混币服务提供商进行协商,约定用户用于混币的输入地址、输出地址、服务提供商的接受地址、返回地址、混币金额、混币输入输出时间、混币手续费等相关参数.

2. 输入阶段:用户按照协商阶段商定的相关参数在约定时间之前将约定资产从输入地址发送到服务提供商指定的接受地址.

3. 输出阶段:服务提供商在约定时间之前将扣除手续费后的资产通过返回地址发送到用户指定的输出地址.

4. 结束阶段:若协议正常运行结束,服务提供商和用户销毁协商阶段留下的记录,保护用户隐私.



## 去中心化混币

尽管中心化混币技术中系列协议在一定程度上保障了资产安全性和隐私性,但其依赖的混币服务提供商仍会带来一些潜在风险,例如遭受黑客攻击进而盗窃用户资产.研究者提出了一系列去中心化的混币协议,通过多方参与的协议代替中心化混币服务提供商,使得用户不需要先将自己的比特币发送给混币服务提供商,而是在网络中找到其他需要混币的用户,通过多方参与者运行协议的方式构造一致的混币交易,确认后签名使得交易生效.这一系列协议从根本上解决了中心化混币存在的信任问题,同时节省了混币服务提供商收取的手续费,但是也存在着一些不足,比如寻找其他混币用户存在困难、容易让外部攻击者混入并监听混币关系甚至进行拒绝服务攻击导致混币失败.去中心化混币协议分为协商、混淆、确认及结束等4个阶段.与中心化混币协议的区别主要在于执行的角色由中心化混币服务器转变为参与混币的用户多方共同完成.

1.协商阶段:用户寻找参与混币的其他同伴,协商去中心化混币协议需要的参数,例如各用户混币输入输出地址、混币金额等参数.

2.混淆阶段:参与混币的用户根据协议对所有输出地址进行混淆,隐藏用户输入、输出地址之间的关联关系.

3.确认阶段:混币用户根据混淆阶段得到混淆后的交易输出构造混币交易,确定无误后进行广播,将混币资产发送到各用户指定的输出地址.

4.结束阶段:若混币协议正常结束,则参与混币的各用户销毁此次混币过程相关记录;若过程出现错误中止,则参与混币的用户找出并且排除造成错误的用户. 去中心化混币技术根据参与方的数量主要分为多方混币技术与双方混币技术两类.其中,

* 多方混币技术参与者数量大于等于3,且隐私保护程度与参与者数量成正相关.其优点在于多参与方增强了地址混淆的外部隐私性,多参与方构造一笔交易也能节省交易费;缺点在于参与者数量的上升会增大攻击者混入的概率,攻击者可以在协议过程中监听并分析其他参与者的输入输出地址关联关系,威胁内部隐私性,甚至进行拒绝服务攻击中断协议进程.
* 双方混币技术只有2个参与方,单次混淆只能提供一定的外部隐私性,并且不能提供内部隐私性,因此,双方混币技术中,用户需要与不同参与方进行多轮混币来增强隐私性.优点在于提升了协议的隐私性,并且每次的混币操作简单;缺陷在于多次混币需要进行多次交易,带来了高额交易费.

# **信息隐藏机制**

地址混淆机制能够在一定程度上保护账本隐私,但是地址混淆的结果仍会在公开账本中存储,攻击者可以通过分析带有特征的混淆交易,在一定程度上威胁用户隐私.为了增强隐私性,研究者们尝试将区块链账本中记录的信息,包括交易发起者、交易接受者、交易金额等进行加密隐藏.另一方面,攻击者可以直接在区块链网络中监听节点信息以及通信情况,将链上内容与节点的真实情况关联,威胁用户网络隐私,可以通过隐藏网络节点信息和节点间流量来保护网络隐私.

账本信息隐藏主要防范区块链账本公开带来的账本分析攻击,通过对账本中隐私数据进行加密,保护账本隐私.并且通过密码学技术提供“凭证”,保持区块链账本正确性的可验证.现有的账本信息隐藏机制的实现技术大多属于零知识证明技术[2],零知识证明即除了所讨论命题的正确性之外,不传达任何其他知识的证明.零知识证明协议中存在两种角色——证明者和验证者:证明者向验证者提供证据证明某命题成立,但不泄露除结论外的其他任何信息;验证者需要验证该证明的正确性.

因此,零知识证明协议需要满足3个重要属性.

(1)完整性:如果证明者能够证明命题成立,那么他能提供证明说服验证者.

(2)可靠性:如果证明者不能证明该命题成立,那么他不能伪造证明欺骗验证者.

(3)零知识性:在验证者验证证明的过程中,不能得到除了命题正确性以外的任何信息.

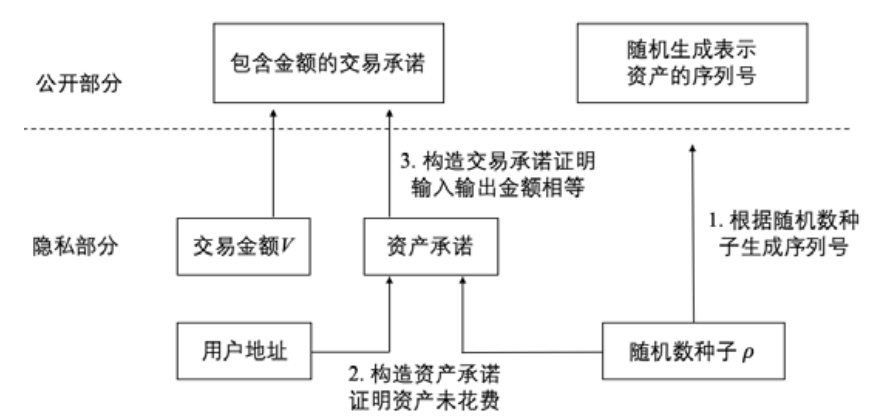
简明非交互式零知识证明技术在非交互式零知识证明证明技术的基础之上进行优化,保持非交互性的同时,减少了证明大小,节省存储空间与验证时间.该技术主要应用于可验证计算领域.非交互式零知识证明技术主要由密钥生成算法、证明生成算法和证明验证算法组成.其中,密钥生成算法由初始化参数生成证明密钥 pk和验证密钥vk,证明生成使用证明密钥 pk 生成对声明 x 和证据 a 的证明π,证明验证算法使用验证密钥验证声明 x 的证明π是否正确.该技术满足以下特征.

• 完备性:对任意正确的声明 x,持有证据 a 的证明者能生成正确的证明π,使验证者验证其正确性.

• 简明性:正确生成的证明π空间复杂度为常数,证明验证算法的时间复杂度仅与声明 x 的长度成线性相关.

• 不可伪造性:不知道证据 a 的证明者无法生成正确的证明π.

• 完美零知识性:验证者除了判断声明 x 正确性以外,不能获取任何信息.



# **通道隔离机制**

通道隔离机制从网络层面对数据进行隔离,保护数据只对通道内节点可见.通过对账本进行隔离,每个节点只处理并存储自己所在通道的数据,防止攻击者访问数据,保护用户隐私.但是通道机制也存在一定缺陷,主要体现在区块链网络中通道部署存在一定代价,节点创建和进出通道需要进行网络配置的修改,灵活性较弱.根据被隔离数据的存放位置,通道隔离机制的实现技术可以分为链下通道隔离和多链通道隔离两大类.

• 链下通道隔离主要应用于高频小额交易,用户通过在区块链上记录起始的状态创建通道,随后在链下进行交易,具体数据通过合约保证安全,但不公布记录在区块链上,需要中止交易的时候,再将最新的结束状态公布并记录在区块链上,终止通道并销毁历史交易记录.

• 多链通道隔离通过在特定节点之间构建独立通信网络作为通道,该网络中信息单独存放在子账本中,非通道内节点不能访问,同一节点可以加入多条不同通道中.多链通道隔离通过在网络层面构建子网络,实现节点通信隔离,杜绝攻击者访问隐私信息,保护用户隐私.

## 链下通道隔离

为了解决区块链账本容量有限的问题,研究者尝试将小额高频交易放在链下的微支付通道进行,仅仅将通道启动和结束的信息记录在区块链账本上.典型的微支付通道技术为比特币系统中的闪电网络技术与以太坊系统中的雷电网络技术.其中,闪电网络技术针对基于未花费交易输出模型的密码货币;雷电网络主要针对基于账户余额状态模型的密码货币,更多利用链上合约机制.该技术主要分为两步:首先,在两个节点之间构造链下双方支付通道;然后,通过节点间的双方支付通道构建支付网络.

• 双方支付通道:在两个地址间构建可信的链下支付通道.通过将资产托管到链上合约创建支付通道,随后参与双方通过对状态更新进行签名确认进行交易,具体过程不需要记录到区块链账本中.当某一方希望中断通道时,将最新状态发布到区块链中,并赎回最新状态对应资产.

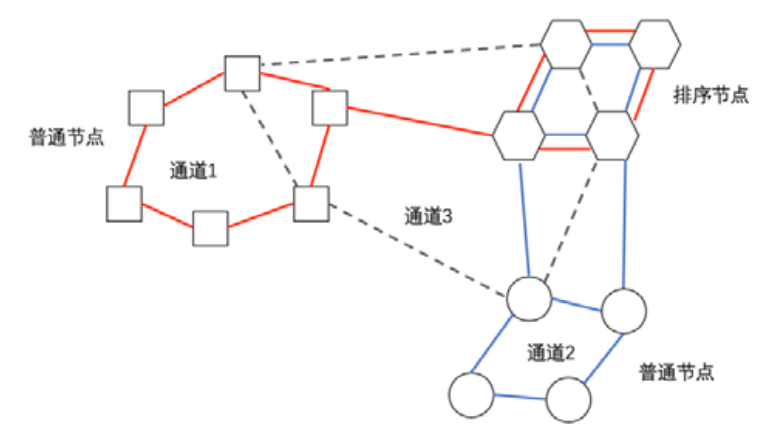
• 构建支付网络:在所有参与用户之间两两构建双方支付通道会带来巨大的存储资源和资产的浪费,因此在双方支付通道的基础上,用户通过已有的双方支付通道进行支付,从而构建全体用户之间的支付

网络.

## 4.2 多链通道隔离

多链通道隔离技术在同一区块链系统中维护多个区块链账本,其中,不同节点群体维护特定子区块链,通过设置访问控制机制保障子区块链数据隐私安全.

目前较为成熟的多链通道技术主要为 HyperLedger Fabric 项目中的通道技术,通道技术主要通过在不同团队的内部节点间分别独立构建区块链,保护内部数据的隐私安全.分片技术将同一区块链账本拆分为多个分片进行维护,各分片只负责维护不同的账户信息以及事务数据,不需要验证全局的账本信息.在必要情况下,分片之间需要进行通信.多通道技术在不同节点间构建互相隔离的通道,维护各自的独立账本,不同通道之间不需要进行通信,因此更强地保护用户隐私数据.



# 总结

随着区块链技术的成熟与众多区块链系统的广泛使用,以及未来区块链技术将在更多的领域中发挥作用,区块链系统中的用户隐私威胁将会成为更加重要的研究问题.本文将区块链隐私内容主要归纳为账本隐私与网络隐私两部分,并分别介绍了所面临的账本分析、网络监听等攻击手段.目前,研究者们提出了一系列区块链隐私保护技术来抵抗这些攻击方式,本文将现有的隐私保护技术归纳总结为地址混淆、信息隐藏、通道隔离这3 大类,并详细介绍了各类隐私保护机制的原理、特征以及不同的实现方式.现有的各类隐私保护机制及实现技术从不同方面保护区块链隐私,因而在实际考虑隐私保护的区块链系统中,通常综合多种技术达到更全面的隐私保护效果.

**参考文献**

[1] Nakamoto S.Bitcoin:A peer-to-peer electronic cash system. 2009. https://bitcoin.org/bitcoin.pdf.

[2] Goldwasser S, Micali S, Rackoff C. The knowledge complexity of interactive proof systems. SIAM Journal on Computing, 1989, 18(1):186−208.