区块链技术及其安全问题

程丽辰 刘吉强
北京交通大学智能交通数据安全与隐私保护技术北京市重点实验室 北京 100044

摘 要 区块链技术具有去中心化、去信任化、透明公开、不可否认等特点，被广泛应用于加密货币的设计与实现。同时，这些良好的特性也可以被用于解决其他领域的发展难题。文章首先阐述区块链的基本概念，并对现有的区块链关键技术进行概述。然后通过分析现有区块链应用中暴露出的问题和技术本身存在的缺陷，探讨其中存在的安全问题。最后总结区块链技术的应用范围、发展前景及有待解决的问题。

关键词 区块链；加密货币；共识机制；隐私保护；哈希函数

引言

随着比特币市场的迅速发展，加密货币及其使用到的区块链技术逐步成为各领域研究和尝试应用的热点。作为未来金融、财政等领域的重要应用技术，区块链技术的安全问题日趋显著。随着计算能力的大幅提升，现有区块链技术所依赖的算力安全将会面临越来越大的风险。而隐藏在其应用领域背后的巨大经济利益，将会使区块链技术受到越来越多的安全挑战，因此，区块链技术的研究正成为国内外的研究热点。Linux基金会于2015年发起超级账本(Hyperledger)项目，旨在搭建区块链技术研究的开放平台，推进该领域的研究。我国也成立了中关村区块链产业联盟、中国分布式总账基础协议联盟、金融区块链联盟、中国区块链研究联盟、陆家嘴区块链金融发展联盟等多个区块链联盟组织，共同推进区块链技术在各领域的研究与发展。作为未来金融等领域构建的重要方法，研究区块链技术的安全性，具有十分重要的意义。

1 区块链关键技术

区块链技术尚未有明确的定义。袁勇等人[1]将区块链技术定义为去中心化基础架构与分布式计算范式，并指出其具有去中心化、时序数据、集体维护、可编程和安全可信等特点。而谢辉等人[2]描述区块链技术为分布式、去中心化、去信任化的技术方案，是一种新的数据分布式存储技术。一般认为，区块链技术包含区块链结构和其他用于数据记录、处理、证明的分布式技术以及密码学技术。

1.1 区块链(Blockchain)

区块链是一条由数据区块有序组成的数据链条，每一个数据区块包含一个时间戳和一个指向上一数据区块的标记。由数据区块组成数据链条的思想在1991年就被提出，但并未引起广泛关注。2009年，中本聪[3]将区块链用于比特币的构建之中。随着比特币的普及，从比特币中提炼而出的区块链及其相关技术得到了工业界和学术界的广泛关注。

1.2 BFT(Byzantine Fault Tolerance)算法

Pease等人[4]探讨了容错系统在故障存在数不同的情况下达成一致的可能性，定义了交互一致性(interactive consistency)概念，并证明当故障节点数大于等于节点总数的三分之一时，不可能达成一个期望的

共识。随后Lamport等人[5]以拜占庭将军问题(Byzantine Generals Problem)对上述进行了更加形象的描述，并讨论了若干种可以达成一致的情形。

现有的拜占庭系统主要分为状态机拜占庭系统和Quorum拜占庭系统[6]。状态机拜占庭系统要求所有服务器以相同的顺序执行请求，并最终使得系统共同维护同一个状态。而Quorum拜占庭系统对请求的执行顺序要求不严格，但要求系统对请求做到响应迅速。

在部分数字货币或其他应用了区块链技术的平台或系统中，使用BFT算法作为共识机制。其中，PBFT算法[7]是BFT算法中较为高效的设计，可以容忍少于系统节点总数三分之一的拜占庭故障节点数量，而且减少了通信开销。Schwartz等人[8]提出了Ripple协议共识算法(Ripple Protocol Consensus Algorithm，RPCA)，可以容忍少于节点总数五分之一的拜占庭故障节点数量。仅当拜占庭故障节点超过节点总数的五分之四时，欺骗交易才会被接收。应用了这种共识算法的Ripple协议，可以在秒级时间内完成一次共识[9]。

1.3 Merkle树(Merkle Hash Tree)

Merkle[10]首次提出这种机制时，将这种树结构用作认证数字签名的一种解决方案，称为tree signature。并且，类似的哈希树结构被大量应用于公钥分配[11]。由于其具有较低的存储开销，可以使用其进行便捷高效的校验，中本聪[3]首次将这种树结构引入数字货币的构建之中，用于建立交易记录的哈希值索引和证明交易在区块中的位置。Merkle树是一棵由哈希值作为节点的二叉树，每一个叶节点的哈希值对应着原始数据的一个片段。父节点的哈希值由左右孩子节点的哈希值计算而来，并以这种方式一直计算到根节点。某小片段在原始数据中的校验过程，就相当于证明小片段对应的叶节点是否在已知根节点的哈希树中存在。

1.4 工作量证明(Proof of Work,PoW)算法

中本聪[3]在比特币中将工作量证明算法用作共识机制，节点对自行选定的交易集合进行竞争性计算，以率先计算出一个低于目标哈希值大小的哈希值。计算出满足要求的哈希值的节点，将计算出的哈希值和与之相关交易集合组成数据区块，并广播给其他节点。其他节点对这一数据区块进行校验，如交易是否合法、哈希值是否正确计算等。如果校验通过，则节点将该数据区块加入区块链，达成共识。区块链不产生分叉依赖于目标哈希值的计算难度，目前的计算难度设计为全网十分钟产生一个区块。

1.5 零知识证明(Zero-knowledge Proof)

Goldwasser等人[12]在上世纪八十年代首次提出了零知识证明的思想，并将其定义为一种不需要提供除论断的正确性以外的任何知识的证明方法。零知识证明中最经典的问题是洞穴模型，A需要在不向B透露开门方法的前提下让B相信A知道洞穴开门的方法。在区块链技术和一些与之相关的应用中，将非交互式零知识证明作为保证匿名性的方法之一。

2 存在的安全问题

区块链技术所具有的去中心化、去信任化、分布式、数据透明的特点，固然为金融系统构建乃至各个领域的发展提供了一种崭新的解决思路。但是，区块链的这些优良的特性，也存在一定的安全问题。

2.1 共识算法的攻击

在去中心化的区块链技术中，每个有意愿维护区块链的网络节点都拥有一份完整的区块链账本备份，并由网络中的共识节点执行共识算法来共同记录账本。而网络中正确的区块链账本，即为大多数网络节点所维护的账本。

在工作量证明算法中，存在51%算力攻击。如果某一个节点或由一些节点组成的组织掌握了全网超过51%的算力，他们就有能力将目前正在工作的区块链转移到另一条包含有恶意行为的区块链上，并使得全网节点在这条恶意的区块链上继续工作。由于比特币所使用的工作量证明算法的安全性依赖于其所消耗的巨大的算力，

51%算力攻击曾一度被认为是难以达到的。然而随着矿池的出现，一个名为GHash.IO的矿池就曾经在2014年6月拥有全网51%的算力；因此，51%算力攻击的威胁始终存在，并且是可能发生的。

在权益证明算法中，存在有“无利害关系”攻击问题(Nothing at stake)。根据权益证明算法，在发生分叉时，节点将在这些分叉上投票，以决定哪一条分叉成为主链。但是，进行投票的节点可能因为私下接收贿赂等原因，做出影响原本判断的决定，从而造成选择分叉上的偏向。

在其他解决拜占庭问题的BFT算法中，共识节点做出正确的判断与网络中恶意节点的数目相关。目前，运行RPCA的服务器多数由Ripple实验室控制。如果由外界运行的服务器数量增加，一旦拜占庭故障节点超过阈值，则有可能阻碍Ripple网络的交易进行，甚至是被迫接受恶意交易。

2.2 区块链上的隐私保护

在以比特币为代表的众多应用了区块链技术的数字货币中，区块链对于网络节点是透明的。即，任何一个节点都可以获得区块链上的所有信息。为保护用户的隐私，比特币中使用由随机数和非对称加密算法生成的地址来代替用户的真实身份。这些又被称为假名的地址，看似能够隐藏用户的真实身份，但只要这些地址直接或间接地与真实世界发生联系，就会失去其匿名性。

例如，Alice拥有2个比特币地址分别是addr1和addr2，每个地址中有1BTC。由于某种原因，Alice想将这些零钱合并为一个地址。则Alice将创造一笔支付给自己的交易，交易输出的地址为addr3。这笔交易被记入区块链之后，Alice又希望通过网络中的某电商平台购买一本价值为2BTC的书籍。那么这将创造一笔由addr3到商家地址的交易，并被记入区块链。此时，这笔支付给商家的交易就和Alice在电商平台留下的包含有真实个人信息(姓名、地址、电话等)的订单信息发生了联系。如果可以同时获得Alice在电商平台留下的订单信息和区块链中的交易信息，就可以分析出addr3的拥有者是Alice。此外，通过在区块链上追溯addr3的交易链，就可以判断出addr1和addr2的拥有者与Alice存在某种关系。

基于类似的思想，Reid等人[13]使用上下文分析和流分析(context discovery and flow analysis)等技术对比特币网络不同时段的拓扑结构进行分析，来寻找一个所谓的“比特币大盗”。Ron等人[14]对比特币网络中历史交易数据的分析，发现了一些交易拓扑结构图中的静态属性，总结了用户在比特币交易中所共有的交易特点，并分离出网络中的巨额交易。

鉴于此，为实现更好的匿名性，Miers等人在[15]中提出了Zerocoin。Zerocoin在比特币网络上进行了扩展，增加了铸币的过程。其实质是将用户之间的比特币进行混淆，使得对比特币的交易拓扑结构图分析是不可能的。用户的铸币过程，就是将自己的比特币兑换成Zerocoin的过程。当用户需要花费这些比特币时，再将比特币赎回，即将持有的Zerocoin重新兑换为比特币。为保证匿名性，比特币兑换前后的信息不能有任何关联。为在交易中隐藏交易双方的身份，Zerocoin使用双重离散对数证明实现对币的所有权证明。但是双重离散对数会带来很大的计算开销，因此，Ben-Sasson [16]建议使用zk-SNARKs来代替现有的证明方式，Danezis等人[17]证明，使用zk-SNARKs可以明显降低Zerocoin的验证开销。

无论是原始的Zerocoin协议还是之后的改进版本，其在比特币网络中的扩展应用可能会为诸如洗钱一类的经济犯罪提供便捷，而且还存在着币值固定、用户之间不能够通过Zerocoin进行直接交易等缺点；因此，Ben-Sasson等人[18]提出了Zerocash，其不仅支持直接交易任意价值的币，而且将交易信息大小和认证时间也大幅下降。Zerocash中沿用[16-17]中建议使用的zk-SNARKs作为零知识证明方法，将交易和币放在区块链上，交易双方通过零知识证明花费或取回币。

2.3 哈希碰撞

在现有的区块链结构中，哈希值是保证区块链不可篡改的重要参数。如果可以构造出具有相同哈希值、但内容不同的数据区块，则可以在该数据区块上发生篡改。如果有足够多的节点共同实施篡改，其他节点将很难判断哪一个数据区块是正确的。

在比特币论坛(Bitcoin Talk)上，Todd[19]在2013年9月发布了为SHA-1、SHA-256和RIPEMD160碰撞设置的长期奖金。其中，SHA-1和比特币采用的Git版本控制系统相关，而SHA-256和RIPEMD160用于生成比特币的锁定脚本和解锁脚本。为获得这项奖金，参与者需要提交两个包含有不同数值的消息，但这两条消息的消息摘要相同。这项奖金迄无人领取，但这并不代表比特币中涉及到的哈希函数就安全。

SHA-1算法曾一直被警告存在风险，而在2017年2月23日，荷兰阿姆斯特丹Centrum Wiskunde & Informatica研究所和谷歌公司的研究人员使用改进后的Shattered算法找到并公开了一例SHA-1哈希碰撞实例[20]。这表示SHA-1已不再具有很好的安全性，需要被安全性更高的哈希算法代替。

2.4 编程安全

区块链技术通常依赖数学方法来建立信任关系[2]，在应用过程中需要编写较为复杂的程序。因而在数学逻辑上，甚至是编程所使用的语言上存在的漏洞都将对区块链技术带来威胁。

2010年8月15日，比特币区块链的第74638块上被发现了一条包含有92233720368.54277039 BTC的交易记录，而且这些比特币仅被发送到两个比特币地址[21]。而导致这次攻击的原因，是由于比特币代码中的大整数溢出漏洞。为使这笔交易失效，比特币核心开发人员开发了比特币补丁版本，并启动了硬分叉。在33个区块的竞争之后，带补丁版本的区块链才成为主链，消除了原有漏洞的影响。

3 区块链技术的应用

早在上世纪九十年代，就曾经兴起过一阵电子货币(e-cash)的研究浪潮。但是，诸如Brands[22-24]一类的电子货币依然停留在有中心的阶段，即账目由一个类似于银行的机构进行管理。而在比特币中，通过区块链、分布式技术、共识机制的应用，由全网节点代替银行的职能。因为网络中的节点之间不需要建立信任关系，这种技术很快就被借鉴到其他种类的加密货币中。迄今为止，类似比特币的加密货币已经达到二百余种。

区块链技术除了在加密货币领域得到了大规模的应用之外，在其他领域也得到了一些应用。

在金融领域，Ripple[25]提出RPCA共识算法，并采用区块链技术搭建了一个为全球银行解决交易问题的平台，实现银行间的即时结算、资金实时追踪，降低了银行间的结算成本。

在密码学领域，Fromknecht等人[26]使用区块链技术对公钥基础设施PKI进行改进，提出了Certcoin。Certcoin将每个实体所拥有的公钥和身份标识发布在公开的区块链上，以达到追踪实体操作和公钥更改的目的。然而，在Certcoin中，实体在PKI上所进行的任何操作都是公开的。为实现隐私保护的目的，Louise等人[27]提出一种隐私可知的基于区块链技术的PKI设计，用户可以根据自己的需求选择是否公开自己的身份信息和曾经使用过的公钥。

在数据处理领域，Healthcare Data Gateways[28]是一个使用区块链技术设计的医疗健康数据共享架构，患者可以轻松拥有、控制和分享自己的医疗健康数据。在保证患者隐私的同时，使医疗健康系统更智能。并且引入安全多方计算，使不可信第三方能够对医疗健康数据进行处理，又不破坏对患者隐私的保护。而MedRec[29]将矿工的角色赋予给现实中的研究人员、公共健康管理者等，让矿工维护数据和网络的奖励等价于对医疗数据

的整合和匿名化。但对于是否公开个人数据的权利，仍然掌握在患者或数据提供者手中。Zyskind等人[30]则利用区块链技术的分布式特点，构建分布式个人数据保护系统，对用户个人数据进行保护。这个系统将区块链技术用作自动访问控制管理，来代替可信第三方。

更综合的，以太坊[31]是一个利用区块链技术和P2P网络来维护的一个可编程平台，开发人员可以借助这一平台开发具有交易、担保、公证等功能的系统，从而将区块链技术应用在更多的领域。

4 发展前景

随着区块链技术的不断发展，其应用领域也越来越广阔。Swan[32]探讨了区块链的发展蓝图，提出了区块链1.0、区块链2.0和区块链3.0的概念。其中，区块链1.0是指加密货币或与其相关的数字支付系统，区块链2.0包括在链上实现的应用于金融、市场、财政等领域的智能合约。而区块链3.0的概念则是区块链在其他更为广阔的领域中的应用，例如管理、医疗健康、基础服务等。

与工业界对区块链技术的热捧相比，学术界对区块链的研究相对滞后。尤其是在区块链协议本身的安全性上，缺乏更深入的研究。在区块链技术中应用的共识算法方面，现有的共识机制或者存在计算量大的问题，或者存在容忍恶意节点数有限的问题。这就需要设计新型的共识机制，使得区块链技术在效率和安全性上都得到很好的提升。

现有的许多区块链技术及其应用，并未对区块链的访问控制进行很好的设计。然而，在区块链3.0时代，链上数据的公开将是可控的，例如医疗数据、政务信息等；因此，在区块链技术中加入访问控制策略就显得尤为重要。将区块链技术与现有的访问控制策略进行结合，也是需要被迫切关注的问题。

5 结束语

区块链技术作为一种新兴的应用在加密货币领域的技术，正在得到学术界和工业界越来越多的关注。如何将区块链技术在其他领域进行拓展，如何在区块链中应用更加高效、更加安全的技术，是需要进一步研究的问题。本文简要介绍了区块链技术的相关概念和基本技术，分析并总结了现有区块链技术中可能存在的安全问题，以期为区块链技术的研究发展提供一些积极的启发。

参考文献

[1] 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

[2] 谢辉,王健.区块链技术及其应用研究[J].信息网络安全,2016,(9):192-195

[3] Satoshi N.Bitcoin: A peer-to-peer electronic cash system[J]. Consulted, 2009:1-9

[4] Pease M,Shostak R, Lamport L.Reaching Agreement in the Presence of Faults[J].Journal of the ACM, 1980, 27(2):228-234

[5] Lamport L,Shostak R,Pease M.The Byzantine Generals Problem[J].Acm Transactions on Programming Languages & Systems,1982,4(3):382-401

[6] 范捷, 易乐天, 舒继武. 拜占庭系统技术研究综述[J]. 软件学报, 2013(6):1346-1360

[7] Miguel O T D C. Practical Byzantine Fault Tolerance[J]. ACM Transactions on Computer Systems, 2001, 20(4):398-461

[8] Schwartz D, Youngs N, Britto A, et al. The Ripple Protocol Consensus Algorithm[EB/OL]. https://ripple.com/

[9] Armknecht F, Karame G O, Mandal A, et al. Ripple: Overview and Outlook[C]//Trust and Trustworthy Computing. 2015:163-180

[10] Merkle R C. A Certified Digital Signature[M]//Advances in Cryptology — CRYPTO' 89 Proceedings.Springer New York,1989:218-238

[11] Merkle R C. Protocols for Public Key Cryptosystems[C]// Security and Privacy, 1980 IEEE Symposium on. IEEE, 1980:122-122

[12] Goldwasser S, Micali S, Rackoff C. The knowledge complexity of interactive proof-systems[J]. Siam Journal on Computing, 1989, 18(1):186-208

[13] Reid F, Harrigan M. An Analysis of Anonymity in the Bitcoin System[C]// IEEE Third International Conference on Privacy, Security, Risk and Trust. IEEE, 2012:1318-1326

[14] Ron D, Shamir A. Quantitative Analysis of the Full Bitcoin Transaction Graph[M]// Financial Cryptography and Data Security. Springer Berlin Heidelberg, 2012:6-24

[15] Miers I,Garman C,Green M,et al.Zerocoin:Anonymous Distributed E-Cash from Bitcoin[C]//Security and Privacy. IEEE, 2013:397-411

[16] Ben-Sasson E. Universal and affordable computational integrity [EB/OL].[2017-04-01].https://www.youtube.com/watch?v=CjUNj8ow6UE&list=PL8hxk6hqV_vkcsn0JSn-LXXpQuChHsG8c

[17] Danezis G, Fournet C, Kohlweiss M, et al. Pinocchio coin: building zerocoin from a succinct pairing-based proof system[J].Workshop on Language Support for Privacy Enhancing Technologies, 2013(1):27-30

[18] Ben-Sasson E, Chiesa A, Garman C, et al. Zerocash: Decentralized Anonymous Payments from Bitcoin[C]// IEEE Symposium on Security and Privacy. IEEE Computer Society, 2014:459-474

[19] Todd P. BitcoinTalk[EB/OL].[2017-04-01].https://bitcointalk.org/

[20] Stevens M, Bursztein E, Karpman P, et al. The first collision for full SHA-1[EB/OL].[2017-04-01].http://marc-stevens.nl/research/papers/SBKAM17-SHAttered.pdf

[21] Jgarzik. Strange block 74638[EB/OL].[2017-04-01]. https://bitcointalk.org/index.php?topic=822.0

[22] Brands S. Untraceable off-line cash in wallet with observers[C]//Annual International Cryptology Conference. Springer Berlin Heidelberg, 1993: 302-318

[23] Brands S. An efficient off-line electronic cash system based on the representation problem[J/OL].[2017-04-01]. https://courses.csail.mit.edu/6.857/2010/handouts/brands-ecash.pdf

[24] Koblitz N, Menezes A J. Cryptocash, cryptocurrencies, and cryptocontracts[J].Designs,Codes and Cryptography, 2016, 78(1):87-102

[25] Ripple Protocol Consensus Algorithm[EB/OL].[2017-04-01].https://ripple.com/

[26] Fromknecht C, Velicanu D, Yakoubov S. CertCoin: A namecoin based decentralized authentication system[EB/OL].[2017-04-01]. http://courses.csail.mit.edu/6.857/2014/files/19-fromknecht-velicann-yakoubov-certcoin.pdf

[27] Axon L. Privacy-awareness in blockchain-based PKI[J/OL].[2017-04-01]. https://ora.ox.ac.uk/objects/uuid:f8377b69-599b-4cae-8df0-f0cded53e63b

[28] Xiao Y, Wang H, Jin D, et al. Healthcare Data Gateways: Found Healthcare Intelligence on Blockchain with Novel Privacy Risk Control[J]. Journal of Medical Systems, 2016, 40(10):218

[29] Azaria A, Ekblaw A, Vieira T, et al. MedRec: Using Blockchain for Medical Data Access and Permission Management[C]// International Conference on Open and Big Data. IEEE Computer Society, 2016:25-30

[30] Zyskind G, Nathan O, Pentland A. Decentralizing Privacy: Using Blockchain to Protect Personal Data[C]// Security and Privacy Workshops. IEEE, 2015:180-184

[31] Ethereum[EB/OL].[2017-04-01].https://www.ethereum.org/

[32] Swan M. Blockchain: Blueprint for a New Economy[M]. O'Reilly Media, Inc. 2015

作者简历

程丽辰

博士研究生，主要研究方向为隐私保护、区块链安全。

刘吉强

博士，教授，博士生导师，主要研究方向为可信计算、隐私保护等。

Blockchain Technology and Security Architecture

Cheng Lichen
Liu Jiqiang
Beijing Key Laboratory of Security and Privacy in Intelligent Transportation, Beijing Jiaotong University, Beijing 100044, China

Abstract Blockchain technology has been widely applied in the cryptocurrency due to its decentralized, trustless, transparent and undeniable features. Meanwhile, all of these sound features can also be used in other fields to solve some difficulties in development. Firstly, this paper explaines the basic conception of blockchain, and introduces an overview of the existing key techniques in the blockchain. Secondly, this paper discusses its potential security problems by analyzing exposed problems of existing applications and vulnerabilities of the technology. Finally, this paper summarizes the application scopes, development prospect and future problems of the blockchain technology.

Keywords Blockchain; Cryptocurrency; Consensus; Private Preserving; Hash Function

(上接29页)

Research on Security Technologies for Home Internet

Liu Minghui
Wang Xiaodi
Gong Xue
China Unicom Research Institute, Beijing 100176, China

Abstract Home Internet brings people convenient, comfortable life. At the same time, privacy leaks, equipment intrusion and other information security issues emerge. In this paper, through the analysis of the typical family Internet crack cases, the security risk points faced by the home Internet are given and the corresponding security technology scheme is put forward for the development of the family Internet business. In the home Internet business system development, testing, operation and maintenance and other aspects, strictly safety technologies and standards should be used to reduce security risks. Regular safety penetration testing should be implemented to identify security problems, and preventive measures be taken to solve the security problems.

Keywords Home Internet; Smart Device; Security