计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.210600152

基于区块链的去中心化众包技术综述

李　玉[1]　段宏岳[1]　殷昱煜[1]　高洪皓[2]
1 杭州电子科技大学计算机学院　杭州 310018
2 上海大学计算机工程与科学学院　上海 200444
(liyucomp@hdu.edu.cn)

摘　要　区块链技术可以广泛应用于各种服务,如在线微支付、供应链跟踪、医疗记录共享以及众包。将该技术应用到众包系统中,可以得到一个去中心化的、隐私保护的、可验证和可追溯的众包服务平台。随着区块链技术的发展,出现了许多基于区块链的众包解决方案,但是缺乏对相关研究的综述。目前研究人员主要从两个角度对去中心化的众包系统展开研究:基于智能合约的去中心化众包平台、基于区块链架构的去中心化众包平台。文中详细综述了主要的基于区块链的去中心化众包的相关工作,并且总结了已有技术中出现的问题,如区块链系统的安全性、智能合约的安全性以及隐私保护的相关问题,并对这些问题展开了详细讨论。最后展望了该领域未来的可研究问题,并提供了大量的可参考文献。

关键词:区块链;众包;隐私保护;智能合约
中图法分类号　TP399

Survey of Crowdsourcing Applications in Blockchain Systems

LI Yu[1],DUAN Hong-yue[1],YIN Yu-yu[1] and GAO Hong-hao[2]
1 School of Computer Science,Hangzhou Dianzi University,Hangzhou 310018,China
2 School of Computer Engineering and Science,Shanghai University,Shanghai 200444,China

Abstract　Blockchain technology can be extensively applied in diverse services,ranging from online micro-payment,supply chain tracking,digital forensics,health-care record sharing to insurance payment. Extending the technology to the crowdsourcing,we can obtain a verifiable and traceable crowdsourcing systems. Emerging research in crowdsourcing applications exploits blockchain technology to optimize the task assignment and reward payment using various consensus protocols and blockchain techniques,which can provide additional security,automatic,verifiable and traceable crowdsourcing platforms. In this paper,we conduct a systematic survey of the key components of crowdsourcing blockchain and compare a number of popular blockchain applications. In particular,we first give an architecture overview of popular crowdsourcing-blockchain systems by analyzing their network structures and protocols. Then,we discuss variant consensus protocols for crowdsourcing blockchains,and make comparisons among different consensus algorithms. Finally,we look forward to the future research problems in this field,and provide a large number of references.

Keywords　Blockchain,Crowdsourcing,Privacy protection,Consensus protocols

1　引言

区块链技术目前最广泛、最主要的应用就是数字加密货币系统,它对现代的中心化金融体系产生了巨大的冲击。如果说比特币[1]提出了一个去中心化的金融系统,那么以太坊[2]等平台则通过对智能合约的支持,为区块链技术带来了更多的可能性。通过区块链技术以及在区块链平台上部署的智能合约,大量的区块链应用被开发部署。区块链技术被应用扩展到了很多领域,如供应链溯源、医疗数据共享以及众包。

众包自2006年被Howe[3]提出以来就受到了广泛的关注。众包作为一种新的商业模式,使得人们可利用互联网上的众包平台分配任务、寻求创意或解决技术问题。传统的众包应用普遍基于一种中心化的结构。中心化的众包模型不管

到稿日期:2021-06-21　返修日期:2021-08-20
基金项目:国家重点研发计划(2020YFB2103805);国家自然科学基金(61802098,61802093);浙江省自然科学基金(LY21F020018);浙江省高校基本科研业务费专项资金(GK199900299012-025)
This work was supported by the National Key Research and Development Project (2020YFB2103805),National Natural Science Foundation of China (61802098,61802093),Natural Science Foundation of Zhejiang Province (LY21F020018) and Fundamental Research Funds for the Provincial Universities of Zhejiang (GK199900299012-025).
通信作者:殷昱煜(yinyuyu@hdu.edu.cn)

是在开发应用程序还是系统管理方面都具有简单性这一优势。常见的中心化众包应用有 Upwork[4]，Amazon's Mechanical Turk[5]等。

然而，众包系统在蓬勃发展的同时，也受制于传统的中心化系统架构。首先，单点故障是中心化系统不可避免的缺陷，这造成了可用性方面的风险。其次，众包系统的数据库中往往保存了大量的用户隐私信息（如姓名、邮箱、电话号码等），这造成了用户隐私信息泄露的风险。然后，当众包系统中的请求者和工人之间产生矛盾时，用户依赖中心化众包平台给出了一个主观的仲裁，这造成了公平性方面的风险。最后，中心化众包平台会收取一定的服务费来维持平台的运营与盈利，这也会降低用户参与众包的热情。

将众包系统部署到区块链上，可以一定程度地解决上述问题。首先，通过区块链技术将众包系统去中心化，可以很好地解决单点故障的问题，有效地提高了系统的可用性。其次，区块链技术为用户提供基础的匿名服务，一定程度上保护了用户的隐私。此外，区块链技术可以保证在没有第三方的情况下提供一个雇主与工人直接对话的平台。同时智能合约技术保证了雇主和工人的雇佣契约无法被单方面撕毁，强调了公平性。最后，由于没有第三方平台的介入，雇主和工人能够节省一定的用于支付给中心化众包平台的费用。可以说，区块链提供了一个实用的平台来解决众包应用由于中心化架构产生的问题。

但是，在区块链上部署众包应用仍然具有挑战性。首先，众包系统从传统的中心化应用转变成基于区块链的去中心化应用，系统也会面临新的安全性问题，如区块链的安全性问题以及智能合约的安全性问题。其次，众包的工作流程在离开了中心化平台的管理之后，将如何适应中心化的区块链环境。然后，区块链的匿名性虽然一定程度上保护了用户的隐私，但是链上暴露的问题依然存在，如何为用户提供额外的隐私保护也是区块链众包系统需要考虑的问题。最后，由于智能合约的执行通常需要消耗计算资源，区块链系统也会对计算资源的消耗进行收费，如何优化智能合约在链上消耗的成本也是重要的问题。

文献[6]讨论了区块链技术对现实生活中众包任务的用例带来的好处，主要分析了区块链在数据库方面为众包系统带来的好处，但缺乏对已有的基于区块链众包系统的总结和分析。本文综述了目前的去中心化众包任务的相关研究。目前研究人员主要从两个角度对去中心化的众包任务展开研究：基于智能合约的去中心化众包解决方案和基于区块链架构的去中心化众包解决方案。针对区块链的安全性、隐私保护等问题，本文讨论了主流的解决方案。对于区块链高资源消耗的特性，本文从合约的优化方面总结了减少去中心化众包成本的最新研究成果。

本文第 2 节介绍了区块链和众包系统的背景知识，总体概述了基于区块链的众包系统；第 3 节详细介绍了众包和区块链的系统架构，综述了基于区块链的众包系统的相关研究；第 4 节分析讨论了基于区块链的众包所面临的问题及其解决方案；最后总结全文。

2　背景知识

2.1　区块链

2.1.1　区块链性质及其技术特点

区块链是一种由块与块链接而成的数据结构，是一个有时序的块列表。区块的结构如图 1 所示，大体上可以分为区块头和区块体两个部分。区块体包含一个交易列表，交易与交易之间通过哈希指针链接成一棵二叉树，被称为 Merkle Tree，这样的结构保证了交易列表数据的不可篡改性。区块头中包含 Merkle Tree 的二叉树根节点、上一个区块的哈希值以及该区块自己的哈希值。区块通过包含上一个区块的哈希值，将自己与上一个区块“链接起来”。这样的区块与区块之间通过哈希指针而构成的链就称为区块链，哈希指针保证了区块链的区块顺序和数据都无法被篡改。

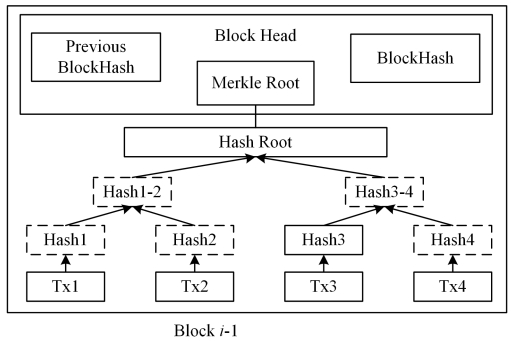

图 1　区块链示意图

Fig. 1　Diagram of blockchain

区块链技术主要的特点包括以下几个方面。

（1）去中心化：区块链系统的一切过程包括数据生成、验证、存储、传播以及维护，这些过程都是在去中心化的系统结构上完成的。区块链底层是一个对等式网络（Peer-to-Peer network，P2P 网络），网络中的每一个节点都具有相同的权力与义务。同时，节点与节点之间通过分布式共识算法对区块链系统状态达成一致。去中心化带来的分布式特性也使得区块链系统具有高可用性，极大地避免了单点故障的问题。

（2）数据不可篡改与可追溯：区块链系统通过哈希链的方式完成区块与区块之间的链接。由于哈希算法的不可碰撞的性质，一旦修改某一个区块的历史交易，就需要对整个区块链系统的哈希链进行修改。同时，区块链网络中的所有参与节点都拥有一致的数据副本，因此单个恶意节点对区块链数据的修改是无效的。并且通过区块链保存的历史区块，我们可以追溯任意一个数据被写入区块链的历史行为。

（3）部分的隐私保护：区块链系统通常采用非对称密码体系中的公钥来标识用户的身份。用户在完成交易时只需要公开自己的公钥或者由该公钥生成的地址，而不需要公开自己的真实身份。特别地，每个用户可以拥有多个地址，来保证自己的匿名性。但这种匿名性仍然不是完全的隐私保护，文献[7]详细地论证了比特币系统中存在的隐私问题，并为比特币的安全隐私问题提出了未来的研究方向。

下文通过一个转账的用例来说明区块链系统的工作流程。如图 2 所示，A 发起一笔转账交易，然后 A 通过自己的私钥对该交易签名，并将该交易通过 P2P 网络广播到区块链网络中。区块链网络中的节点在接收到该交易后会将其转发给其他的节点，同时对该交易进行验证。如果该交易通过验证，则会被打包到下一个将要生成的区块中。最后在节点与节点之间通过分布式共识算法来决定获得下一轮记账权的节点。获得记账权的节点将新的区块发布广播到全网中的所有节点，节点收到区块后进行验证与同步。此时，该交易被写入区块链中，转账交易完成。在比特币系统中，通常会等待 6 个新区块的产生，才会确认一笔交易是安全的成交状态。

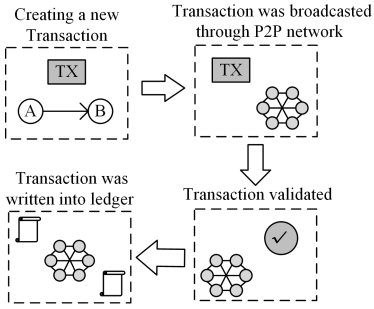

图 2 区块链的工作流程

Fig. 2 Workflow of blockchain system

2.1.2 公有链和私链

区块链技术可以按照不同的去中心化程度进行分类[8-9]，大体上可以分为公链和私链。公链又称公有区块链，它具有完全的去中心化性质，如比特币、以太坊这一类区块链系统，它们允许所有人参与，任何节点都可以随时参与整个区块链系统的维护。每个节点的地位平等，节点之间的共识基于 PoW 算法或者其他共识算法而非基于信任。

私链又称联盟链，通常应用于企业级的系统。相比公有链，联盟链最显著的特点就是权限控制。与公有无许可区块链网络不同，联盟链中的参与者节点彼此都是已知的。虽然参与者可能不完全信任彼此，但网络可以在参与者之间确实存在的信任的治理模式下运行，如处理纠纷的法律协议或框架。同时，Hyperledger Fabric[10] 平台支持可拔插的共识协议，根据不同的场景切换相应的共识协议。例如，当其部署在一个企业中，或者由一个可信的权威机构操作时，完全拜占庭式的容错共识可能被认为是不必要的，并且会对性能和吞吐量造成过度的拖累。在这种情况下，崩溃容错(CFT)共识协议可能已经足够。而在多方、分散的用例中，可能需要更传统的拜占庭容错(BFT、多层 PBFT[11])共识协议。Quorum[12] 也是常见的联盟链平台。

由于公链具有更好的流通性以及普及性，基于区块链的众包系统通常部署在公链上。但联盟链可以提供完善的权限管理和更有力的隐私保护，这对于众包系统而言也是十分重要的属性。最后，公链与联盟链的对比结果如表 1 所列。

表 1 公有链与私有链的对比

Table 1 Comparison of public blockchian and private blockchain

属性	公有链	联盟链
节点数量	不限制	有限制
去中心化程度	完全去中心化	多中心化或者弱中心化
激励机制	代币激励	无激励机制
节点准入机制	节点自由进出	节点必须经过审核才准入
匿名性	用户身份对系统匿名	用户身份需经过审核
共识机制	安全性高但效率低的 PoW，PoS 等	安全性始终效率较高的 CFT，BFT 等

2.1.3 区块链上的智能合约

智能合约这一概念最早由 Szabo[13] 于 1995 年提出，他将智能合约定义为："一个智能合约是一套以数字形式定义的承诺，包括合约参与方可以在上面执行这些承诺的协议。"数字形式表示智能合约必须是计算机可读的代码，也就是说智能合约其实就是一段可执行的程序。

虽然智能合约早在 1995 年就被提出，但是真正得到发展与应用还是在区块链技术得到广泛应用之后。区块链系统为智能合约提供了真正意义上的实用的执行环境。一般来说，区块链智能合约可以定义为：通过一段可执行程序促进、验证和执行区块链上两个或多个双方之间订立的合约的计算机协议。由于智能合约通常在区块链上部署，因此它们具有一些独特的特性。首先，智能合约的程序代码将在区块链上被记录和验证，从而使合约具有不可篡改性。其次，智能合约的执行是在匿名的、无信任的独立节点之间执行的，而没有集中控制和第三方当局的协调，保证了公平执行。最后，智能合约可以有自己的数字加密货币或其他数字资产，并在触发预定义条件时完成对资产的转移[14-15]。

从有条件的资产转移角度来说，任何比特币交易都是一个"智能合约"。比特币交易需要提供有效的加密签名，加密签名被认证之后就可以完成资产的转移。但是由于比特币脚本语言的限制，不可能实现具有复杂逻辑的智能合约。以太坊利用以太坊虚拟机(EVM)实现了支持图灵完备的智能合约的公有区块链平台。整个以太坊系统可以看作一个单例的分布式状态机，以太坊网络中的每个节点都运行一个 EVM 实例，并通过共识机制保证本地 EVM 与其他节点的 EVM 具有相同的状态。许多高级编程语言都可用于编写以太坊智能合约，如 Solidity 和 Serpent，并且合约代码被编译成 EVM 字节码，并部署在区块链上进行执行。以太坊是目前最受欢迎的智能合约开发平台，可用于设计各种去中心化区块链应用程序(DApps)，如数字版权管理、数据共享和众筹等[16]。

智能合约的运作机理如图 3 所示。在以太坊中，智能合约以可执行代码的形式填入一笔区块链交易的数据负载中，然后经过 P2P 网络传播到各个节点中，经过验证打包等流程写入到区块中。区块链系统将自动识别智能合约的可执行代码，并将其部署到区块链系统中。智能合约部署完成之后，以太坊会为该合约创建一个合约账户，同时将可执行字节码存入该合约账户的存储空间中。

智能合约中封装了预定义的若干状态、执行逻辑以及触发条件。以太坊中所有智能合约的执行，都是由外部账户发送交易来触发的。外部账户会在交易中告知以太坊系统需要

调用的智能合约的函数，并且给出输入。当以太坊网络中的节点接收到该交易时，会按照交易中给出的信息执行相应的智能合约，并将执行产生的状态更新后写入区块链。

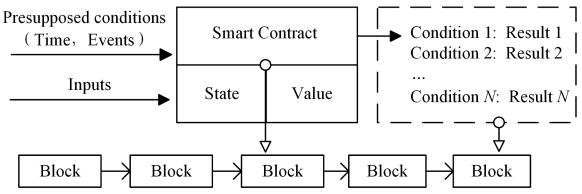

图 3　智能合约的运行机制

Fig. 3　Operating mechanism of smart contract

2.2　众包系统

2.2.1　众包系统的总体介绍

众包(Crowdsourcing)于 2006 年被提出，之后许多学者对众包的定义展开了讨论。文献[17]总结对比了许多工作，给出了众包的基本特征。

(1)采用公开的方式召集大众。

(2)众包任务通常是计算机难以单独处理的问题。

(3)大众通过协作或者独立的方式完成任务。

(4)众包是一种分布式的问题解决机制。

总的来说，众包是一种开放的面向互联网大众的分布式的问题解决机制，它通过整合计算机和互联网上具有相应技能的群众来完成计算机单独难以完成的任务。

2.2.2　应用

众包技术在许多领域中都得到了应用。如在科研领域，Snow 等[18]最先使用众包技术来解决一些机器学习和数据挖掘问题，他们用 Amazon's Mechanical Turk 平台通过众包技术解决了一系列语言注释任务，例如情感识别、文字相似性等计算机难以解决而人类容易解决的问题。在生活中，滴滴打车等网约车平台很大程度上满足了人们的出行需求，同时提供了许多的就业岗位。还有许多地图软件，例如高德地图、百度地图等，每天使用地图软件的个人、出租车司机等都在不停地产生数据并传输给其公司，而这种数据来源也是一种众包。同时地图软件开发人员通过众包得来的数据来优化他们的软件。可以说我们正在经历一个众包的时代，每个人都或多或少地参与了众包。

2.2.3　中心化众包平台的缺点

在上文介绍的众包流程中，请求者和工人之间所有的交互都通过一个中心化的众包平台来进行。本节将总结中心化众包平台容易遭受的攻击类型。

(1)用户隐私问题：传统众包平台通常会收集用户的个人信息，并将其存储在中心化的平台中。而这样的存储了用户隐私信息的平台很容易遭到黑客的攻击，从而导致用户隐私的泄露。

(2)单点故障问题：中心化平台容易产生单点故障的问题。而众包平台就是众包服务中的单点，一旦众包平台遭受网络攻击导致宕机，就会导致众包服务处于长时间不可用的状态。

(3)公平性问题：公平性问题主要分为两种，分别针对请求者和工人。请求者可能虚假报告任务的状态来逃脱对任务奖金的支付，我们称之为雇主欺骗。工人可能会套取众包任务的奖金而拒绝为众包任务做出贡献，我们称之为工人欺骗。在传统的中心化众包平台中，这类问题通常由平台裁定，很容易出现一些公平性的问题。

2.3　区块链＋众包

区块链＋众包的基本思路是使用区块链平台替换传统众包系统中的中心化众包平台，将区块链的技术优势引入到众包系统中。区块链技术的引入对众包系统带来的好处如下。

(1)解决了单点故障问题：区块链是一个可用性很高的平台。截至 2021 年 5 月，etherscan 上显示以太坊的全网全节点数量为 5431。面对如此庞大的分布式系统，企图通过网络攻击使得以太坊全网节点宕机几乎是不可能的。使用去中心化的区块链平台取代传统众包系统的中心化的众包平台，可以使基于区块链的众包系统不再有单点故障的问题。

(2)一定程度上的隐私保护：由于区块链的匿名性，用户在使用区块链平台时不需要提供自己的身份信息。尽管区块链上的历史交易数据都是公开可见的，但是其中并不会包含用户的个人身份信息。但这种隐私保护不是完全的，众包任务中有些可能会涉及工人的位置坐标信息，而这些信息通常被认为是隐私信息。

(3)公平性问题：公平性问题在基于区块链的众包系统中仍是一个重要的议题。相比中心化的众包平台，基于区块链的众包系统更能杜绝主观因素给公平性问题带来的影响。因为裁决过程通常是基于智能合约的自动化过程，而合约内容是任务请求者和工人在任务进行之前商定好的。区块链平台提供的智能合约的可信执行环境能保证智能合约的正确执行，只要合约没有漏洞，通常就能保护合约双方的权益。

(4)更好的开放性：区块链技术的引入使得让更多人参与到众包成为可能，从而提高了任务完成的质量。文献[19]提出了一种基于区块链的众包陪审团系统，用于提高司法质量。对于众筹项目，基于区块链的解决方案在项目规模以及开放程度上也更具优势[20]。

2.3.1　基于区块链的众包解决方案

基于区块链的众包解决方案主要分为两种。一种是基于现有区块链平台，通过智能合约将众包的业务在区块链上实现。其代表性工作是 CrowdBC，该系统主要分为 3 层：应用层、区块链层和存储层。区块链层是保存事务属性的地方。存储层则包括工人完成的任务的详细信息和内容。而应用程序层实现了业务逻辑，提供了用户管理、任务管理和智能合约编译器的功能。

另一种方案则不使用现有的区块链技术，其主要思路是设计并构建一种定制的区块链平台，使其能更好地满足众包任务的要求。代表性的工作是 MCS-Chain，该系统提出了一个基于区块链的移动众包系统，对系统底层区块链平台的共识机制、激励方式以及节点架构做出了深度定制。该系统构建了一个去中心化的移动众包系统，满足了众包用户对任务发布、任务接受、任务质量评估以及奖励发放等功能的需求。

2.3.2　去中心化的众包应用

基于区块链的众包应用也逐渐引起了开发者和学者的兴趣。文献[21-22]提出了一种基于区块链的众包方案，用于解

决众筹这一特定众包问题。

2016 年以太坊上出现了一个很受关注的众包项目，The DAO(Decentralized Autonomous Organization)，意为去中心化的自治组织。我们可以认为 The DAO 是一个风险投资基金，其中的资金由以太坊众筹而来，同时每个参与者的行为以及众筹资金的流动都由预定义的智能合约主导。用户向该智能合约发送以太币来换取代币，以获得审查项目和投票表决的权力。所有的代币持有者都可以对投资议案进行投票表决，一旦某议案获得足够的票数支持，系统就会将相应的款项划给该投资项目。投资项目的收益则会按照一定的规则回馈给众筹的参与者。可以说 The DAO 不仅是资金上的众包，也是协作上的众包、治理上的众包。“The DAO”项目发起众筹，在当时是很伟大的尝试，受到了很大的关注度。其 1 个月内就筹到了 1.5 亿的以太币，众筹的速度和规模前所未有。但是由于重大的智能合约安全漏洞，最终“The DAO”仅仅存活了 3 个月就迎来了失败的结局，甚至导致了以太坊主链中 ETH 和 ETC 的分裂。

尽管“The DAO”项目以失败告终，但是它作为基于区块链的众包应用，引起了人们极大的关注，从这点上来说基于区块链的众包应用仍然具有相当大的价值。

3 系统架构

3.1 众包系统的体系结构

3.1.1 众包工作流概述

传统的众包参与主体包括：任务请求者(Requester)、解决任务的工人(Worker)、中心化众包平台(Platform)。

如图 4 所示，任务请求者和工人通过中心化的众包平台联系在一起。任务请求者(Requester)使用众包系统来完成他的需求时，需要以下几个步骤。

(1)设计任务：任务请求者根据自己的需求，将任务的内容、任务的要求、任务的目标以及完成的方式进行预定义，最后结合市场价格以及任务难度设定合适的激励机制。一个清晰并且激励合理的任务往往更容易得到高质量的结果。

(2)发布任务：将设计好的任务发布到众包平台上，等待工人提交答案。

(3)任务结果质量评估：任务请求者审核评估工人提交的答案，选择接受与否。

(4)整合答案：整合所有工人提交的答案，得出任务的最终结果。

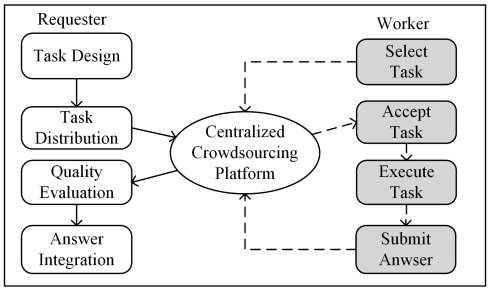

图 4 众包工作流程

Fig. 4 Workflow of crowdsourcing

而对于工人来说，使用众包所包含的步骤如下。

(1)选择任务：工人从众包平台上浏览、查看与自己专业技能相契合以及感兴趣的任务。

(2)接受任务：工人选择合适的任务并接受任务。

(3)执行任务：工人按照任务描述执行任务。

(4)提交答案：工人将自己得出的任务答案提交到众包平台。

可以看出，无论是对任务请求者还是工人而言，他们的直接交互对象都是众包平台，众包平台在整个流程中充当了一个可信的第三方角色。按照时间顺序，文献[17]将众包流程分为 3 个阶段：任务准备阶段、任务执行阶段、任务答案整合阶段。

3.1.2 任务准备阶段

任务准备阶段包括请求者设计任务、发布任务以及任务分配。设计任务是众包服务中最重要的一环。一个合理的众包任务设计应该包含以下内容。

(1)任务内容设计：相比复杂任务，众包更适合用于完成微型简单任务。因此，在利用众包解决大型复杂任务时，应该将大型任务进行分解。

(2)任务激励机制：如何为一个众包任务定价是一个热点问题。过高的定价可能会吸引不诚信工人，从而导致工人欺骗行为，同时也会增加众包任务的成本。而过低的定价会使得任务的吸引力降低，导致任务的完成时间被推迟。众包任务的定价需要从市场关系的角度来考虑，合适的定价是提高任务完成质量的关键因素。

(3)欺骗行为处理：工人有潜在的欺骗，存在任务奖励的可能性，如何检测并惩罚这种行为是提高任务完成质量的关键环节。常见的方法是在众包问题中增加一些已经知道答案的测试问题，如果工人错误地回答了这些测试问题，则可以认为该工人为欺骗者，任务请求者可以不接受该工人的答案。

设计好众包任务之后，请求者通过众包平台发布众包任务，平台按照一定的策略进行众包任务分配。

在实际的众包服务中，任务和工人之间的选择往往是双向的，工人选择任务，任务也选择工人。这一环节通常是由众包平台主导的，工人通过平台搜索感兴趣的任务，平台也给工人推荐合适的任务。任务搜索与任务推荐也是众包领域研究的热点问题。

所谓任务选择工人，指的就是任务推荐。众包平台依据工人的专业领域、兴趣爱好主动地推送任务信息给工人。Ambati 等[23]提出使用工人完成任务的历史信息作为任务推荐的依据。随着人工智能与大数据的发展，基于工人的用户画像的推荐系统逐渐成为了主流的研究方向。

3.1.3 任务执行阶段

任务执行阶段指工人从接受任务、解答任务到提交答案的过程。由于众包任务的种类十分繁多，用户执行任务的过程通常是不可一概而论的。如优衣库自 2005 年以来每年举办全球 T 恤设计竞赛(UTGP)，2019 年举办的第十四届 UTGP 就受到了来自世界各地的 18000 多个参赛作品。对于这种 T 恤设计众包任务，工人只需要在任务的截止日期之前按照任务的要求提交自己的设计稿即可。对于这种比较宽松的众包任务，任务执行阶段的约束就仅有时间约束以及主题约

束。但是类似于滴滴打车这类时空众包任务，任务执行阶段的约束会严格许多，司机（工人）的到达时间、行进路径等都受到了平台的严格约束。

通常来说，任务的执行方法是在任务设计时由任务请求者事先规定好，工人按照任务内容的描述执行任务。工人在这一环节的自由度并不会特别高，决定该环节完成质量的因素主要是任务请求者一开始的设计以及工人自身的专业素质。

3.1.4　任务答案整合阶段

任务答案整合阶段是众包流程的最后一步，如何处理工人们提交的答案是这一阶段的关键问题。由于众包任务的开放性，接受任务的工人的教育背景、专业素养往往参差不齐。面对工人提交的良莠不齐的答案，任务请求者需要对答案进行进一步的处理。简而言之，去其糟粕，取其精华。

最简单的整合策略是将一个任务分配给多个工人，然后收集多份答案，并统计收集到的答案，使用简单的少数服从多数的原则选择出最终结果。通常对于简单的图片识别、数据标记等任务，采用这种方式比较有效。但是这种投票方式预先假定了工人的正确率都相同，实际上工人之间的正确率往往有差距，从而导致整合结果有误差。针对这一问题，文献[24]在任务中增加了一些测试题目，用于估计工人答题正确率的先验概率，然后利用贝叶斯理论结合工人的正确率来整合得出最终答案。这种通过增加测验题来估计工人正确率的方法假定了工人的正确率是一个常数，但是工人的正确率往往是变化的。在任务执行过程中，工人对任务的理解会越来越深，从而正确率会呈现上升的趋势。文献[25-26]提出了一种基于 EM（Expectation-Maximization）算法[27]的答案整合方案，这种方案考虑了工人答题正确率的变化。但出于算法时间复杂度的限制，EM 算法不适用于问题较多或者参与工人数量较多的场景。

3.2　区块链的体系结构

如图 5 所示，区块链系统可以分为 6 个层次：数据层、网络层、共识层、激励层、合约层、应用层。

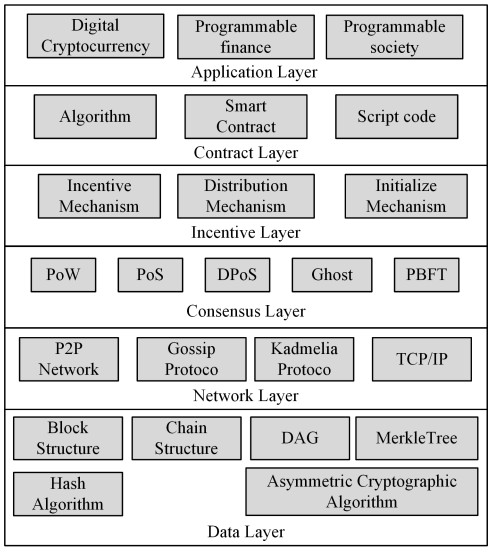

图 5　区块链的系统架构

Fig. 5　Architeture of blockchain

3.2.1　数据层

数据层定义并封装了区块链的底层数据结构，包含了哈希函数、Merkle 树以及链式结构等关键底层过程。

哈希函数具有很多优良的性质，在区块链中几乎无处不在。哈希函数具有单向性，即无法通过输出值逆推输入。并且哈希函数产生的数据的长度都是一致的。哈希函数的输出表现出了一种随机性，即便输入很接近，经过哈希函数之后的输出仍然有显著的区别。最重要的是哈希函数具有不可碰撞性，这里指只要输入值不同，经过哈希函数之后的输出也绝对不同，无法人为地制造产生相同输出的不同输入。

Merkle 树是区块内部一种重要的数据结构，区块链系统使用 Merkle 树来组织区块内部的交易数据。如图 1 所示，Merkle 树包含了区块体的所有交易和一个由交易列表通过哈希运算生成的根哈希值（Merkle Root）。Merkle 树运算过程一般是将区块内的交易数据进行分组哈希，并将生成的新哈希值插入到 Merkle 树中，如此递归直到只剩最后一个根哈希值并记为区块头的 Merkle Root。它可以快速校验区块内部的数据的存在性，并且提供轻量级的零知识证明查询。通过 Merkle 树，客户端在包含 N 个交易的区块中确认一笔交易的算法复杂度可以降低到 $\log 2N$，从底层支持了只保存区块链头信息的轻量级区块链客户端。

区块链系统中最典型的是长链式结构的区块链，如比特币和以太坊。基于有向无环图结构的区块链是一项创新的技术，在微支付以及物联网领域有很大的应用价值，如 IOTA[28] 和 ByteBalls[29]。

3.2.2　网络层

网络层定义并封装了区块链的组网方式和数据传播协议。公有链的组网方式通常采用对等式网络，节点与节点之间权利平等，没有任何特权节点和管理员节点，体现了去中心化的特点。私链以及联盟链的组网方式通常比较复杂，由于其进出网络需要经过许可的特点，通常包含组织管理者等管理节点来管理网络中节点的权限。

区块链系统需要通过节点之间的通信来完成对系统状态的共识。分发新区块、转发新交易等过程都需要数据传播协议来管理，如比特币系统采用的 Gossip[30] 协议和以太坊采用的 Kadmelia[31] 协议。这些协议规定了区块链网络中节点的通信方式，保证了消息传播的高效性和规范性。

3.2.3　共识层

共识层规定了区块链系统中节点之间的共识机制，保证了节点之间的数据一致性。共识机制确保了区块链中的数据正确性和完整性。在公开区块链项目中，PoW（工作量证明）共识机制是应用最广泛的。PoW 鼓励节点对特定的复杂计算难题进行求解，最先求得答案者即竞争得到记账权，成为区块发布者并获得奖励。这种工作量证明机制保证了系统的安全性，但对于系统的吞吐量来说是不小的负担，同时也带来了资源浪费的问题。为了减少资源浪费，提出了 PoS 共识机制。不同于 PoW 的算力竞争，PoS 是一种股份竞争。简单来说，拥有系统权益越大的节点获得记账权的概率就越大。如果说 PoW 机制的安全性建立在数学上，那么 PoS 机制的安全性则建立在博弈论上。PoS 共识机制的关键在于构建适当

的博弈论模型以及相应的验证算法。但是,这种简单 PoS 势必会产生马太效应,使得系统的公平性以及去中心化变得脆弱。而 PBFT[32]算法具有高吞吐量、高可用性,同时对于不可信环境也有一定的容错率。但是其由于通信开销较大,无法适应多节点的公有链环境,因此更多地被应用于节点数量适中且节点可信度较高的私链、联盟链中。

3.2.4 激励层

对于区块链系统而言,激励机制是保证其长久运行的重要手段。通常激励层以发放代币的方式分配奖励,发放给参与区块链共识的节点,用于弥补节点在区块链中记账和校验消耗的成本。“挖矿”指节点之间竞争记账权以获取激励的行为。除了“挖矿”竞争出块奖励,通常每笔交易都会包含支付给节点的服务费。服务费越高的交易就越容易被优先处理。

3.2.5 合约层

合约层包含了区块链系统中各种各样的脚本代码、算法以及智能合约,我们可以将合约层看作控制和管理区块链基础设施的算法和逻辑,它是区块链系统实现可编程性的基础。以太坊支持图灵完备的编程语言实现智能合约,用户可基于以太坊构建带有复杂业务逻辑和循环的去中心化应用,为区块链应用打下了基础。

3.2.6 应用层

应用层将具体的区块链应用程序部署在区块链上,通过合约层中的智能合约与区块链产生交互。利用区块链去中心化的基础设施构建的应用被称为 DApp(Decentralized Application)。DApp 与传统的 App 在结构上区别不大,如果传统 App 由前端和后端构成,那么 DApp 则在前端和后端的基础上再多一个合约端,对于一些简单的 DApp 而言,可以直接由合约端取代后端。可以认为,DApp 是后台代码运行在区块链网络中的 App,因此 DApp 继承了区块链的一些优点。

3.3 区块链+众包的现有工作

使用区块链平台代替传统众包中的集中式众包平台,能给众包技术带来新的生命力。其详细的优点在第 2 节中已经阐述,本节将主要讨论如何在区块链的框架上实现众包的流程。我们讨论总结得出两种类型的解决方案:基于智能合约的解决方案以及基于区块链架构的解决方案。

3.3.1 基于智能合约的解决方案的系统框架

随着区块链技术的发展,以以太坊(Ethereum)为代表的区块链 2.0 平台都支持使用图灵完备的编程语言来编写智能合约。许多研究人员使用智能合约技术将众包系统部署到区块链平台上,以利用区块链平台的各种优点改良众包流程。

CrowdBC[33]提出了一个比较全面的基于区块链的去中心化的众包系统框架,如图 6 所示。

CrowdBC 提出的系统架构分为 3 层:应用层、区块链层和存储层。应用程序层和区块链层属于逻辑层,存储层属于数据层。具有特殊技能的工人可以查询和竞争应用层中由请求者发布的任务。区块链层使用任务状态(待接受、已接受、已完成)作为输入,以实现工人和请求者之间的共识。然后系统将任务元数据(如数据大小、所有者、哈希值、指针)放在区块链层,将原始数据放在存储层。由于区块链的特性,在区块链上进行数据的存储是十分昂贵的,将主要数据存储在存储层可以有效减小数据存储带来的开销。CrowdBC 将众包任务流程的 3 个阶段拆分成了 5 个子过程。传统众包流程中的任务准备阶段被划分成用户注册、任务发布以及任务分配 3 个子过程,而众包任务的执行阶段通常是在链外进行的。最后将任务答案整合阶段划分成奖励发放与任务质量评估两个子过程。

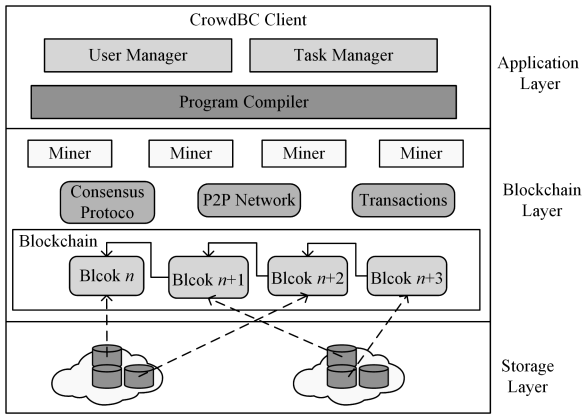

图 6 CrowdBC 的系统架构

Fig. 6 Architecture of CrowdBC

CrowdBC 使用智能合约技术,将 5 个子流程部署到了区块链上。任务发布、任务分配、奖励发放以及任务质量评估的过程都包含在 Requester-Worker Relationship Contract (RWRC)合约中。该智能合约会在请求者发布任务 T 并发布任务信息时创建。当工人想要接收任务时,系统会调用 RWRC 合约中包含的一个验证函数 checkWorkerQualification( )来检查该工人的声誉和可靠性值是否满足最小值。如果工人是合格的,系统则更新 RWRC 合约的状态并将其发送到区块链上,这代表工人成功地接收了一项众包任务。然后工人执行任务并提交答案,工人提交的答案或答案的缩略信息将作为 RWRC 合约中的 solutionEvaluate( )函数的输入,用于众包任务质量评估。整个质量评估过程由区块链中的矿工节点完成,这避免了任务质量评估中的主观性问题。值得注意的是,solutionEvaluate( )函数的编写是由请求者事先预设好的,根据不同的众包任务,评估的方法也有所不同。solutionEvaluate( )函数编写的好坏将直接影响众包任务的公正性。

根据 3.1.4 节介绍的各种众包任务答案整合的方法,简单的投票机制可能很适合该系统的任务质量评估方法,其简单的逻辑使得合约代码编写不至于太复杂,同时计算资源的消耗不算太高,有利于部署在区块链平台上。而使用复杂的 EM 算法的评估方法由于需要占用大量算力资源,并不适用于区块链众包的场景。

3.3.2 基于智能合约的解决方案的相关工作

文献[34]将基于区块链的众包流程细化成了 9 个阶段:初始化、任务提交、任务发布、任务接收、方案提交、方案仲裁、支付、任务回调和服务补偿。特别地,如果在方案仲裁阶段任务发起者和任务接收者没有达成共识,则方案将由整个网络中的专业领域与任务领域一致的工作人员进行投票仲裁。如果领域中有超过一定比例的人群支持该计划,则意味着该解

决方案通过，系统将执行支付费用；否则意味着解决方案失败，任务回滚。文献[35]提出了一种分散的、保护隐私的和基于众包的医疗研究方法。其使用基于区块链的去中心化平台使研究人员能够从大量的志愿者那里寻求明确的帮助。此外，为了避免侵犯隐私和实现相互信任，其使用了零知识证明的密码原语。最后，用智能合约技术以确保患者和研究人员之间公平的互惠交换回报。文献[36]总结并分析了现有区块链众包系统的安全问题：众包任务中固有的隐私问题、链上隐私泄露问题以及智能合约的安全性问题。其提出了安全的BCS(区块链众包系统)的安全性和隐私性要求，并且基于JUICE实现了原型系统，解决了相关问题。文献[37-42]针对不同的问题以及领域，使用智能合约实现了基于区块链的众包系统原型。不同的是，文献[37]的任务分配是基于一种Repeated-Single-Minded Bidder (R-SMB)的拍卖机制来实现的，工人通过对已经发布的任务进行投标来竞争项目。文献[38-39]分别实现了一个简单的区块链众包系统，利用智能合约实现了众包系统的去中心化。整个过程都由智能合约进行约束，这对竞争性的众包任务更加有利。文献[40]针对的是基于区块链的能源众包问题，而文献[41]针对的是基于区块链的软件众包问题，文献[42]使用基于区块链的众包系统将机器学习中的计算密集型任务外包出去。文献[43]分析了现有的区块链＋众包的工作，认为用户隐私在任务分配的环节没有得到可靠的保护，并使用智能合约对任务分配环节进行了优化，在任务分配的过程中保证了用户的匿名性。文献[44]提出了基于区块链的版权保护众包数据交易框架，它可以使相互不信任的参与者能够进行真实和可信的数据交易，同时确保数据的质量和版权，通过基于语义相似度的拍卖机制来选择获胜者，并通过智能合同来确定支付方式，通过语义相似性解决了需求匹配问题，保证了交易的质量。其采用了结合区块链和数字指纹技术的版权保护方案，确保了数据所有者的合法权利。文献[45]提出的区块链众包系统中包含了一个二次投票机制，与文献[34]相似，用于解决系统中的纠纷。文献[46]针对流媒体视屏转码问题，提出了区块链众包方案，设计了一种智能合约可以在视频转码众包过程中同时实现招标功能和任务执行功能。文献[47]与文献[33]相似，给出了一个通用的区块链众包解决方案，其核心功能由3个智能合约实现：Identity contract，Credit contract，Task contract。文献[48]与文献[41]相同，将区块链众包用于软件众包过程，不同的是它更关注层次定价机制与市场稳定性。文献[49]提出了一个针对时空众包任务的区块链众包解决方案。与上文提到的相关工作不同的是，该工作将任务按照特定的规则进行排序，对于高优先级的任务将会优先分配工人去完成，这就要求工人的工作是统一服从系统安排的，只有这样才能实现任务可靠的分配。文献[50]第一个设计并实现了具有私密、匿名性质的区块链众包系统，解决了去中心化众包系统中的数据泄露和身份泄露问题。

表2列出了上述的相关工作，可以看到大部分的区块链众包项目都部署在公链(以太坊)上。由于公开区块链平台的开放性，在公开的区块链平台上部署众包系统在推广方面可以享受到很多优惠。同时由于公开区块链通常有良好的应用接口和软件生态，研究人员在设计智能合约时有更多的工具可以使用。由表2可知，大部分的区块链众包工作都选择在公链上进行部署，但值得注意的是公链和联盟链在智能合约机制上的差别并不大，大部分的公链上的工作经过部分修改之后也能移植到联盟链平台上。

表2　基于智能合约的解决方案的相关工作总结

Table 2　Summary of work related to smart contract-based solutions

	公链	联盟链	额外隐私保护	合约消耗优化	存储离链
文献[33]	√			√	√
文献[34]	√				
文献[35]	√		√		
文献[36]	√		√	√	
文献[37]	√		√		
文献[40]		√			
文献[42]	√				
文献[43]	√		√		
文献[44]	√			√	√
文献[45]	√			√	√
文献[46]	√				
文献[48]	√				
文献[49]	√				

虽然区块链本身的匿名性提供了基础的隐私保护，但由于链上暴露的情况存在，这种简单的匿名机制对用户的隐私并不能起到足够的保护。文献[35]使用零知识证明的密码原语(zk-SNARK)来保护病患的隐私信息。文献[36]使用组签名技术为用户提供了更强的匿名性，同时使用AES(一种典型的对称加密方案)保护了众包任务结果的机密性。文献[37]提出的秘密拍卖机制能保护用户竞标信息的隐私性。文献[43]引入了一个独立的KM(Key Manager)机构，每个用户注册时都必须经过KM认证并初始化。合约的部署也通过KM来完成，通过KM机构为用户提供了一套额外的匿名机制。同时使用可加密搜索技术参与任务分配，将众包任务的具体信息也在区块链上进行加密，切断了用户与众包任务在区块链上显式的联系，一定程度上解决了链上暴露的问题。

除了隐私保护之外，减少智能合约的消耗也是区块链众包系统需要考虑的问题。如表2所列，上文中的相关工作主要使用存储离链技术来减少智能合约在区块链上的数据存储，以达到减少智能合约消耗的目的。如CrowdBC[33]引入了一个分布式数据库来作为存放原始数据的存储层。CPchain[44]提出的基于区块链的版权保护众包数据交易框架采用了“链上交易，存储离链(on-chain trading，off-chain storage)”的机制，将数据存储在独立于区块链的IPFS分布式存储数据库中。文献[45]也将原始数据存储在IPFS上。而文献[36]将计算消耗较高的组签名以及相关的加密原语操作在链外执行，一定程度上降低了智能合约的消耗。

以上的区块链众包系统在合约的优化方面考虑得并不多，使用的方法也比较简单，这可能会造成比较严重的成本问题。关于这个问题也有相关研究者给出了切实可行的方案，本文将在4.3节中详细讨论该问题。

3.3.3　区块链众包系统的智能合约的基本模型

基于区块链的众包系统在工作流程上与传统众包系统差别不大，总体上来说仍然可以划分为本文3.1节中介绍的众包的3个阶段：任务准备阶段、任务执行阶段以及任务答案整

合阶段。而其区别在于区块链众包系统使用智能合约技术在区块链上实现与部署众包工作流程的3个阶段;同时区块链众包系统需要对智能合约进行合理的设计,以保证工人和请求者之间的公平性。本文分析相关工作并总结了一个区块链众包系统中智能合约函数的基本模型。

(1)注册函数

该函数用于新用户的注册,同时存储用户的相关信息。该函数体现了系统对用户的建模,对于简单众包任务,用户包含的信息可能相对较少,而对于复杂的时空众包任务,用户包含的属性可能比较复杂(如位置、技能、时间等)。由于区块链系统本身不适用于存储大量数据,如果用户的数据量较大,则可以考虑只将用户信息的元数据(原始数据的索引或者哈希)存储到区块链中,如CrowdBC中就单独开辟了一个存储层用于存储原始数据。

(2)任务发布函数

该函数用于新用户的注册,同时存储任务的相关信息。与注册函数不同的是,请求者在发布任务的同时应该在合约中存入一定的保证金,以防止请求者不按照要求支付赏金。因此保证金的数额应该至少能够覆盖任务的赏金,保证金机制几乎在所有的区块链众包系统中都有应用。

(3)任务分配函数

任务分配是众包过程中最重要的一个环节,它决定了每个工人将执行的任务。目前大部分的工作都是基于Worker-Select-Task(工人选择任务)模型的,即将区块链平台作为一个任务发布的大白板,工人可以在白板上自行选择感兴趣的任务。对于这类模型,最直观的是文献[37]提出的基于拍卖的任务分配机制。工人们对自己感兴趣的任务发起报价(即自己理想的报酬),然后系统采用贪心策略来完成工人们的竞标过程,最后给出任务分配的结果。

不同的是,文献[49]针对Task-Select-Worker(任务选择工人)模型提出了一种任务分配机制。其将当前不可分配的任务按照发布时间排序和先来先服务(FCFS)原则为任务分配工人。根据工人的位置信息、报价信息以及众包任务的相关信息,采用贪心策略为任务匹配工人。该方法优先考虑了众包任务的分配率,而不是工人的兴趣,而且需要一批随时处于待命状态的工人。其在提高任务分配率的同时,也提高了管理的难度。

(4)任务质量评估函数

该函数主要涉及众包系统的激励机制和评估任务的完成情况,并根据评估结果将赏金发放给工人。该环节是区块链众包系统中的难点。如3.1.4节所述,关于传统众包中的众包任务答案整合的问题,学者们提出了许多方法,但是这些方法并不全都适用于区块链众包的环境。由于区块链环境的特殊性,函数不能设计得太复杂,运行时间应该尽可能短。最简单的策略是将一个任务分配给多个工人,统计收到的答案,使用少数服从多数的方法得出最终结果。同时,给出正确答案的工人将会获得全额赏金,给出错误答案的工人将无法获得全额赏金。对于数据标记、图片识别等任务而言,该方法是有效的。但是对于一些主观性较强的时空众包任务而言(如修理水管),该方法显然不适用。

因此,任务质量评估函数不存在通用的范式,对于不同的众包任务,任务质量评估的方法都不尽相同。当区块链众包系统真正投入使用时,应该考虑绝大多数的众包任务类型,并且按照任务类型使用合理的任务质量评估方法。

3.3.4 基于区块链架构的解决方案框架

上文提到的解决方案大多数集中在区块链的应用层,也有许多研究者在区块链体系结构中的更深层次进行了优化与定制,使其能够更好地适应众包的应用场景。Feng等[51]认为大多数目前流行的区块链项目都面临着效率低下、集中化以及分叉问题。同时,为了解决传统集中式移动众包面临的单点故障等问题,其提出了一个基于区块链的移动众包系统。

系统模型如图7所示,节点与节点之间通过P2P网络互相连接。节点主要分为3类:终端用户、工人、矿工节点。终端用户是整个系统中的消费者,他们是众包任务的发起者,支付劳酬给工人的同时也支付服务费给矿工节点。工人节点根据协议接受并执行众包任务,矿工节点合作维护和管理为移动众包服务的区块链。区块链作为去中心化的移动众包平台,记录移动众包过程,并评估所有系统实体的信任度。在MCS-Chain中,每个矿工节点都保留一个区块链的副本,并可以访问存储在区块链上的数据。下文主要介绍该系统的工作流程。

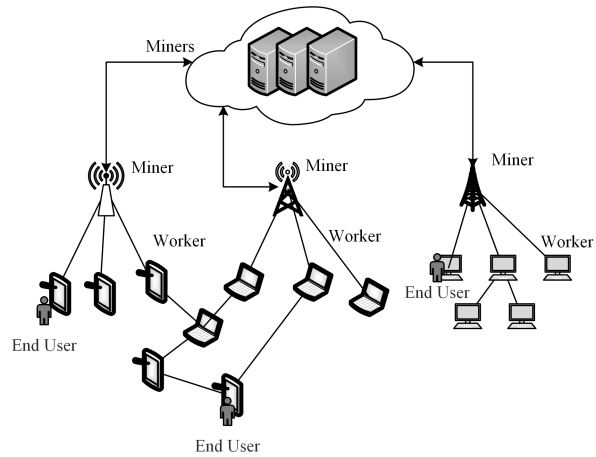

图7 MCS-Chain的系统架构

Fig. 7 Architecture of MCS-Chain

(1)终端用户向P2P网络发送一个众包任务请求消息,该消息包含了终端用户的签名,以及任务的相关信息。矿工节点接收到消息后会验证该消息的签名以及相关信息,如果合法则将该消息记录在区块链上。

(2)任务请求消息被验证之后,工人可以发起报价申请执行该任务。经过几轮的讨论,终端用户与工人就雇佣关系达成共识并向P2P网络广播一条表明雇佣关系且包含双方签名的消息。矿工节点检验消息中的签名以及相关信息的合法性,若合法则将该消息记录在区块链中。

(3)工人执行任务,将答案数据经过加密后上传给终端用户。终端用户按照预先提出的标准评估任务质量,如果合格,则支付一笔钱用于矿工的劳酬和矿工的服务费;如果不合格,终端用户可以拒绝付款或者要求工人重新完成任务。同时评

估结果会上传到 P2P 网络中供矿工节点检查评估结果是否有效，防止不公平的事件发生。

(4)任务完成之后，任务中所有的参与者根据自己的行为对他人的工作生成反馈。矿工节点会在预期的和指定的时间内收集反馈。当反馈收集期满后，他们将进行信任评估和区块生成，以确认任务的执行和支付。

区块链系统中恒定的出块时间限制了系统的吞吐量。不同于流行的区块链系统，该系统设定了一个支付量阈值。一旦矿工节点记录的移动众包行为所产生的总支付超过设定好的阈值，就将所记录的信息打包成一个区块发布出去。同时该系统设计了一种基于节点信誉的共识机制，保证了同时只有一个节点能发布区块，从而避免了区块链的分叉。

3.3.5　基于区块链架构的解决方案的相关工作

文献[52]设计了一个基于区块链的移动众包框架，该框架在保护参与者隐私的同时保持了移动众包服务的完整性。其使用了一些聚类算法对请求者生成服务策略，利用了一些决策算法在最大化服务时间、增加利润和减少供应商的能耗之间实现权衡。文献[53]针对区块链的底层技术，提出了一种改进的信任证明(PoT)共识机制，它正好适用于众包服务场景。改进后的 PoT 共识机制使用主观逻辑算法来优化共识节点的选择，降低了恶意成员参与众包的可能性，同时确保了所有的成员都有平等的机会参与众包活动。文献[54]提出了一种新的雾区块链分布式众包声誉管理方法，以防止无线物联网系统中用户的隐私泄露、恶意用户的参与和声誉篡改。通过基于雾计算的层次结构，灵活地将用户的身份和任务分离开来。此外，针对众包任务的多约束要求，其提出了一种多因素声誉评价方法来准确地识别恶意用户。文献[55]详细介绍了一个系统的设计和实现，该系统将基于区块链的移动众包用于患者与医疗工作者之间安全地交换医疗数据；同时使用基于加密数字货币的方式来奖励对系统做出贡献的用户。文献[56]提出了一种混合区块链来实现去中心化和隐私保护的众包平台。其部署 DPOS 和 PBFT 共识，降低了众包的能源消耗，提高了交易验证率，加快了众包流程在区块链上完成的速度。文献[57]将联邦学习与区块链众包系统相结合，用于智能家居厂商的人工智能模型的训练。用户通过手机收集自己的智能家居数据来训练厂商的初始模型，然后将训练好的模型发送到区块链上。其使用区块链取代了传统联邦学习系统中的中心化聚合器，利用区块链的不可篡改性追溯厂商或者用户的恶意行为，最后使用差分隐私方法来保护用户的隐私，使得厂商无法通过用户上传的训练模型推断出用户的隐私信息。文献[58]认为，基于区块链的去中心化众包系统因为其开放性，必然会带来恶意用户的问题，从而提出了一种基于区块链的去中心化信任管理方法，用于管理众包任务、工人和最终众包结果的可信度，其中通过将已知结果的子任务插入任务来计算可信度；并且开发了一种共识机制 POCW，如果矿工捕获更多的懒惰行为，采矿难度就更小。其基于这种共识机制开发了一个可信度区块链，以方便地管理众包网络的信任。文献[59]分析了区块链平台对众包系统带来的性能瓶颈，并且提出了一个创新的混合区块链众包平台 zkCrowd。zkCrowd 集成了混合区块链结构、智能合约、双账本和双共识协议，以确保通信、验证交易和保护隐私。其利用 DPOS 和 PBFT 的共识协议，显著提高了交易验证效率，降低了众包系统中的交易延迟和能耗，使用零知识证明和智能合约的权限控制在保护用户隐私的同时也能完成对众包任务质量的评估。

表 3 列出了基于区块链架构的解决方案的区块链众包系统的相关工作。下文按照 3.2 节中区块链的层次进行进一步的对比总结。

表 3　基于区块链架构的解决方案的相关工作

Table 3　Summary of the work related to blockchain-based solutions

	数据层	网络层	共识层	激励层	合约层	应用层
文献[51]			√	√		
文献[52]					√	
文献[53]			√	√		
文献[54]	√				√	
文献[56]	√		√			
文献[57]					√	
文献[58]			√			
文献[59]	√		√			

区块链的数据层主要定义了数据的存储结构以及存储过程。区块链由于具有不可篡改的性质，在区块链众包系统中常用来存储众包任务的记录、任务以及用户的相关元数据。文献[54]设计了基于雾计算的跨层隐私保护模型，在数据层上进行了分层。通过分离用户身份与任务信息，减小数据之间的联系性，降低链上暴露的风险，从而达成了保护用户隐私的目标。文献[56]将区块链分为公链和子链，在数据结构上将区块链系统分层，达成了隐私保护的目标。同时，智能合约可以根据应用场景的不同，选择在公链还是子链上进行部署。文献[59]采用了双链、双账本的设计，在这种机制下，系统既能保护用户的隐私，也能权衡数据的透明性。

如表 3 所列，上述工作都没有提及网络层上的工作。相对于网络层的传播协议等内容来说，区块链众包系统的研究重心更加偏向于上层的一些应用。虽然对于区块链系统而言，网络层涉及到系统安全以及共识一致性等关键性问题，但是研究网络层的数据转发传播机制对于区块链众包系统而言并没有太大的意义。

许多工作都在共识层上做了优化改进。由于目前流行的公链项目中大多数使用 PoW 共识机制，这造成了吞吐量低下以及资源浪费的问题。同时 PoW 共识也存在着潜在的中心化以及 51%算力攻击的风险。针对以上问题，文献[51]设计了一种新的共识算法。它保证即使同时出现几个生成的块，也能确定唯一的块；同时对于中心化趋势做了一些优化，使得发布区块较少的节点有更大的可能性获得记账权。文献[53]提出的 PoT 共识机制能够自动完成众包过程中参与方的信誉评估，同时在防止恶意成员的方向上做出了许多优化。文献[56,59]使用 DPOS 和 PBFT 共识来提高系统效率。文献[58]设计了一种新颖的共识机制 PoCW，如果矿工捕捉到更多的懒惰动作，将会更容易获得系统的记账权，也就是挖矿更容易。这种机制鼓励网络中的节点去发现并监督众包流程中的懒惰用户行为。

激励层定义了系统对区块链网络中维护节点的奖励机制。文献[51]除了给予节点经济上的激励,还有声誉奖励。每次众包任务结束后启动的信任评估机制会对服务节点的声誉进行调整,使得用户能够选择可靠的工作者。文献[53]提出了一种基于博弈论的激励机制,通过设定适当的价格,来确保诚实是每个节点的最佳策略。该方法能够鼓励验证节点主动检查并提供可信的验证结果,最大限度地防止节点与恶意节点合谋。

合约层是区块链可编程性的基础,其中定义了许多控制和管理区块链基础设施的算法和逻辑。虽然上述文献很少使用到智能合约的字样,但是只要是管理区块链的算法和逻辑,都可以被认为是合约层的工作。文献[52]在合约层上定义了一种聚类算法和决策算法,提高了众包服务的质量。文献[54]在合约层定义了一种多因素声誉评价方法,该方法考虑了结果的准确性、完整性、任务的完成时效性以及工作人员的历史声誉。文献[57]提出的基于区块链的众包联邦学习系统,要求用户上传由自己的隐私数据训练而成的机器学习模型。为了保护用户的隐私,在合约层定义了差分隐私过程。

在基于区块链架构的解决方案中,研究者们对区块链在不同的层次上进行了改造定制,使得设计的系统拥有更高的吞吐量、更好的性能以及更好的适用性。但是,对比公链项目,这种经过改进的区块链众包系统往往面向的用户人群比较狭窄,这对于众包这项利用人群智能的技术而言是一个不可忽视的缺点。

4 众包+区块链中存在的问题及其解决方案

4.1 安全性分析

4.1.1 区块链安全性

使用区块链平台替换传统集中式众包系统中的中心化众包平台,解决了单点故障以及公平性和隐私性的问题。区块链作为新兴的技术,其开放性和分布式的特点为系统带来了优点,同时也为系统安全带来了风险。为了更好地研究基于区块链的众包系统,本节讨论基于区块链的众包系统面临的安全问题。

4.1.1.1 双花攻击

双花攻击是区块链系统最常见的一种攻击类型,它针对区块链的底层数据结构,因此不管是公链还是联盟链系统,只要是区块链系统,都无法从根本上防止双花攻击。从字面的角度来解读,双花攻击指攻击者通过某种手段将一笔钱花费两次的行为。而实际上,双花攻击指攻击者针对区块链平台的共识机制,改写区块链底层数据,最终将区块链系统中共识的数据改写成攻击者想要写入的数据。它破坏了区块链系统的不可篡改性,对区块链系统的共识一致性是毁灭性的打击。对于区块链众包系统而言,双花攻击可以回滚历史数据。试想,一个工人通过区块链众包系统完成了一个众包任务,整个过程都记录在区块链中。但攻击者通过双花攻击将区块链数据回滚到该工人完成任务前的状态,那么该工人所做的工作将无法被证明有效,因此也无法拿到众包任务的赏金。对于区块链众包系统而言,这是不可接受的糟糕情况。

但是,几乎所有的区块链平台都会对双花攻击进行防范。比特币和以太坊这类公链使用了最长主链原则和工作量证明机制来抵抗双花攻击。企图实施双花攻击的攻击者需要与区块链全网节点竞争算力,以构建一条比主链更长的支链来回溯交易历史。但是,随着主链越来越长,攻击者能够成功构建更长的支链的概率是呈指数下降的。这对于控制算力不超过50%的攻击者来说,双花攻击是几乎无法实施的。联盟链平台通常具有严格的身份审核和准入机制,从拒绝攻击者进入系统的角度防范了双花攻击。而且联盟链通常不采用工作量证明机制,而是采用拜占庭容错的共识机制,这使得双花攻击在联盟链的环境中难以展开。

4.1.1.2 51%算力攻击

对于普遍采用工作量证明的公有区块链系统而言,掌握全网超过50%以上的算力就意味着对区块链系统的完全控制。攻击者可以轻而易举地破坏区块链的安全性和去中心化。

目前全球规模的大矿池机构盛行,这是公有区块链系统面临的重大挑战。联盟链通常不使用工作量证明机制达成共识,而且联盟链通常具备多中心化的性质。故51%算力攻击只在公有区块链中讨论。

51%算力攻击对于上文提到的基于智能合约的区块链众包解决方案来说是巨大的安全问题。对于基于区块链的众包系统来说,一旦所依托的区块链平台遭到破坏,作为区块链上层应用的区块链众包系统的可用性、安全性自然就无法得到保证。

4.1.1.3 粉尘攻击

区块链系统的区块大小的出块时间往往是规范化的,方便管理共识的同时也带来了系统扩展性的问题。粉尘攻击利用区块链系统低吞吐量的漏洞,通过大量广播无意义的垃圾交易来阻塞区块链网络,从而导致大量交易阻塞、无法得到处理的情况。当公链网络中出现大量粉尘攻击时,用户需要提高交易中的交易费以吸引矿工节点的优先打包处理。而联盟链网络中出现大量粉尘攻击时,系统真正想要执行的交易则会被粉尘阻塞,使得系统陷入不可用状态。但联盟链具有权限控制的功能,当发现粉尘攻击的恶意节点时,可以调整相关恶意节点的权限,即将其从联盟链网络中移除。

粉尘攻击也会降低区块链众包系统的任务处理速度。以CrowdBC系统为例,请求者需要调用Requester-Worker Relationship Contract(RWRC)合约去发布一个众包任务,此时请求者会向区块链网络发起一笔合约调用交易。但是,由于区块链网络中的"粉尘"太多,使得该交易无法及时得到处理。因此,系统如果察觉到区块链网络遭到粉尘攻击,则应该调高发布的交易中的交易费,以换取更快的响应。

4.1.1.4 女巫攻击

联盟链平台通常使用拜占庭容错共识机制(PBFT)来保证系统的共识,而PBFT算法只能抵抗不超过节点总数1/3的恶意节点。只要网络中的恶意节点不超过临界值,整个系统的共识就不会被破坏。而女巫攻击指恶意节点利用这个特性,将自己伪装成多个节点,参与拜占庭容错过程,只要恶意

节点复制的身份超过全网节点总数的1/3，就可以破坏整个系统的共识。与4.1.1.2节提到的51%算力攻击不同的是，前者针对的是工作量证明的共识机制，而女巫攻击则针对的是拜占庭容错的共识机制。

联盟链为了防止女巫攻击，可以加强身份认证，使用更加复杂的身份认证机制来增加身份伪造的难度；或者使用容错率更高的共识算法，以增加女巫攻击的难度和成本；也可以通过硬分叉等手段挽回女巫攻击产生的损失。

对于基于联盟链的众包系统而言，女巫攻击一旦成功，后果与上文提到的双花攻击类似，都会使区块链底层的数据发生变化，造成数据回滚，损害用户的权益。

4.1.2　智能合约的安全性

区块链众包系统非常依赖智能合约的设计与执行。3.3.3节给出了区块链众包系统的一个通用的智能合约框架，从设计的角度讨论了智能合约的安全性。本节从代码细节的角度讨论了智能合约的安全性。

由于区块链的不可篡改特性，智能合约一旦部署到区块链上就难以修改。因此，代码一旦出现漏洞，就很容易被不法分子利用，从而导致巨大损失。2016年6月17日The DAO[60]遭到攻击，由于其智能合约中存在一种被称为重入漏洞的错误，攻击者利用该漏洞窃取了价值巨大的以太币。同时，该事件还导致以太坊分化成了两条链，严重破坏了以太坊的社区生态。2018年4月22日，黑客利用美链的智能合约中整数溢出的漏洞，凭空套取了大量的代币。诸如此类的利用智能合约漏洞套取利益的事件层出不穷。

对于区块链众包系统而言，智能合约出现代码漏洞是不可接受的错误。黑客极有可能利用合约漏洞去盗取请求者的保证金，对用户造成无法挽回的损失。文献[61]总结分析了智能合约中常见的代码漏洞。

(1)调用栈深度攻击

以太坊EVM设置的调用栈深度为1024，攻击者迭代调用合约1023次之后再发布交易触发该合约，从而导致合约执行异常。

(2)可重入攻击

在一个合约中调用另一个合约时，当前合约进程会停下等待另一个合约调用结束。攻击者利用这个中间状态，在合约未执行结束时再次调用合约，实施可重入攻击。The DAO事件的攻击者就是通过不断递归调用withdrawblance函数，取出本该被清零的以太坊账户余额，从而窃取了大量以太币。

(3)整数溢出攻击

攻击者利用智能合约编程语言Solidity的整数类型范围的限制，干扰智能合约的执行，构造超出整数范围的变量和中间计算结果，使得程序运行异常并从中获利。

由于区块链的不可篡改性，智能合约一旦部署上链就无法更改，因此智能合约在部署上链之前应该进行详细的安全性检查。文献[62]提出的Oyente可以基于符号执行对智能合约的漏洞进行检测。它以智能合约的字节码和以太网全局状态作为输入，输出合约存在的安全问题。但Oyente作为较早提出的智能合约漏洞检测工具，对于严重的漏洞效果较好，但是对于合约自杀和滥用发送源等方面的问题的检测效果并不理想。文献[63]在Oyente的基础上改进了与整数溢出相关的漏洞检测。文献[64]提出了Securify，它使用安全模式和违反性模式来判断智能合约的安全属性。Securify对智能合约字节码进行反编译，然后分析智能合约以推断语义事实，主要包括数据和控制流的依赖关系。相比Oyente，Securify为智能合约提供了更全面的分析，但是它没有提供可靠性证明，可能会导致一些潜在的安全性问题。Remix提供了一个基于浏览器的智能合约编写和漏洞修补的应用，它使用内嵌的静态分析工具对预定义的漏洞进行检测。文献[65]提出了一个通用函数式编程工具，支持验证工具的自动执行，实现了对智能合约的语义正确性和运行过程安全性进行验证。

但现有的形式化验证和程序分析工具都只能针对已知的智能合约漏洞进行检测。这对智能合约的安全性来说是不够的，未来的研究应该更加关注对智能合约的动态监测的程序分析工具[61,66]。

4.2　隐私保护

用户在区块链中使用公私钥对来标识自己的身份，每个用户可以注册多个公私钥对，同时公私钥中不包含用户的个人身份信息。通过这样的匿名机制，区块链为用户提供了基础的隐私保护。但是区块链数据需要在整个区块链网络中的所有节点之间进行复制，这要求区块链中的数据必须是公开且可验证的。这种透明性是区块链系统很有吸引力的原因，但同时也会造成隐私泄露的风险。面对这样的开放式网络环境，区块链系统面临着链上暴露的风险。分析人员通过分析链上公开的交易记录，可以获得用户的交易规律，甚至能够推测用户的身份信息和位置信息[67]。对于众包应用，众包任务信息中很有可能包含工人或者请求者的敏感身份信息，如位置信息、专业技能、身份信息等。恶意工作人员也可以通过简单地复制和提交他人上传的任务结果作为自己的数据来获得奖励，因为区块的确认不是实时完成的(例如，通常在数据提交给网络后，以太坊需要大约12s来完成确认)，这对区块链众包系统的激励机制是毁灭性的打击。不仅是用户的隐私，任务数据的机密性也是区块链众包系统需要注意的问题。下文介绍一些实用的用于区块链众包系统的技术。

4.2.1　环签名

为了避免联系紧密的交易数据被攻击者利用，许多学者提出了基于零知识证明、环签名等密码学技术的交易混淆机制。如图8所示，环签名将多个签名信息按照一定顺序构造成首尾相连的环，签名者可以用自己的公私钥和其他多个签名者的公钥来生成一个环签名。整个过程中签名者不需要经过其他签名者的同意，只需要知道他们的公钥信息就可以完成环签名的过程。通过环签名，签名者将自己的信息隐藏在环中的同时，也能保证签名的可认证性、不可否认性以及消息完整性。环签名后来发展成了门罗币[68]的核心协议，但是环签名技术的扩展性差、签名长度长的问题也会对区块链中的应用效率产生不良影响。针对这一点，文献[69]提出的基于双线性对运算的环签名方案优化了传统环签名的运算效率。

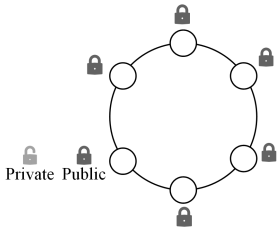

图 8 环签名示意图

Fig. 8 Example of ring signature

4.2.2 零知识证明

零知识证明体系广泛应用于区块链技术中，比特币中的轻节点快速验证交易(SPV)机制就使用了零知识证明机制。零知识证明中有两类参与者：声称某项命题为真的证明者、想要验证命题真实性的验证者。零知识证明要求证明者在不透露任何其他信息的情况下，向验证者证明某一命题的真实性。在区块链众包系统中具体表现为：矿工节点不需要知道众包任务以及任务双方的任何隐私信息也能完成对众包任务的验证过程。零知识证明有效地保护了用户的隐私，其他人仅仅知道发生了有效的交易，但是对交易的具体信息并不了解。2014 年，Sasson 在 Zerocoin[70] 的基础上使用简洁非交互零知识证明(zk-SNARK)[71]，构造了一个匿名支付协议 Zerocash[72]。该协议实现了交易双方身份和交易金额的隐私保护。zk-SNARK 技术具备抗量子攻击的性质，但该技术尚未成熟，存在一定的效率问题。

4.2.3 安全多方计算

除了使用密码学手段对区块链提供隐私保护，文献[73]总结了安全多方计算对区块链隐私保护的支持。

如图 9 所示，安全多方计算可以在排除可信第三方的条件下，协调多个参与方协同计算某一个约定的函数。在整个计算过程中，每个参与方仅能获取自己的计算结果，无法通过交互数据推测其他参与方的输入信息和输出信息。安全多方计算技术提供的隐私保护和去中心化特性与区块链技术的核心思想不谋而合。文献[74]设计了一种部署在区块链上的安全多方计算协议，保证了签署消息的合法性和匿名验证签名的高效可行性。一个切实的联想是，在基于区块链的众包系统 CrowdBC 中，任务质量评估函数 solutionEvaluate( )可以利用安全多方计算协议，让众包任务的参与者都参与到任务质量评估环节。这不仅能增加众包任务答案整合的质量，也进一步保证了公平性。

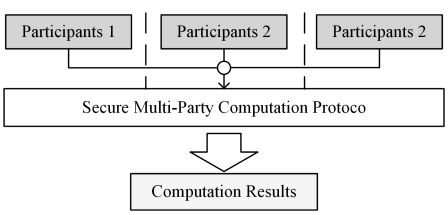

图 9 安全多方计算的示意图

Fig. 9 Example of secure multi-party computation

4.3 合约优化

区块链平台虽然提供了安全的、开放的去中心化服务，但是这种去中心化服务并不是免费的。在以太坊中，任何会引起计算机资源消耗的操作都会产生交易费(gas)，如数据存储、哈希计算等。准确来说，交易费是一种工作量单位，合约的工作量越大，消耗的交易费就越高。交易费机制的存在可以防止一些死循环智能合约的出现，也是维护以太坊系统安全的重要属性。尽管交易费机制对以太坊的安全具有重要意义，但是也导致了现有基于区块链的众包系统面临的成本问题，特别是当众包项目规模很大时。换句话说，中介费不是通过去中心化的众包系统来消除的，过去支付给集中的众包平台的钱，如今要支付给分布式的基础设施，以获得相应的服务。

去中心化众包项目对比传统的集中式众包系统可能没有相当大的经济优势。通过 CrowdBC 完成的众包图像标记任务的花费很可能是集中的 AmazonMechanicalTurk 收取的费用的 4 倍以上。同样，通过去中心化的人群感知系统[75]收集来自 1000 个供应商的数据的费用可达 170 美元。减小众包项目在区块链上的开销，是推动基于区块链的众包系统发展与应用的关键问题。

减少智能合约在区块链上产生的消耗的议题引起了学术界的许多关注，下面介绍一些相关的工作。

为了解决减少智能合约在区块链上产生的消耗问题，文献[76]提出了 GasReducer，完成了对智能合约在字节码层面的优化。该工作在前期分析了以太坊的智能合约工作流，总结得出了 24 种可优化序列。首先对以太坊智能合约字节码进行反编译，然后将得到的汇编代码序列进行分析。对比之前提出的 24 种可优化序列，将每个发现的可优化特征代码序列替换成优化后的更高效的有效代码。

文献[77]提出了一套新的协议，称为 NF-Crowd，它提供了一个可靠的解决方案，将去中心化众包项目的成本与人群的规模脱钩。NF-Crowd 主要从操作离链和任务聚合两个角度讨论了如何减小在区块链上的开销。

(1)操作离链

1)存储离链：NF-Crowd 使用 IPFS[78]来存储原本需要存储在区块链上的大文件。IPFS 是一个分布式的文件系统，它保证了时间戳和文件内容的安全性和不可篡改性。在区块链中只需要保存大文件的 IPFS 链接，即可解决在区块链上的存储成本问题。

2)计算离链：将智能合约的执行环节放到区块链外，以减小相应的链上计算开销。同时，NF-Crowd 通过使用链上计算来备份离链计算，不正确的离链计算结果将永远不会在最终判决中被采用。如果有不诚实的参与者出现，NF-Crowd 则会将计算重新加载到区块链上进行，用链上计算的正确结果替代离链计算的结果。

(2)任务聚合

对大型众包项目的整合也是降低消耗的重要手段之一。NF-Crowd 定义了两种任务类型。$1\times N$ 型：创建的单笔交易的成本随着人群规模的增加而增加；$N\times 1$ 型：创建的交易数量随着人群规模的增加而增加。

$N\times 1$ 型任务聚合策略的核心思想就是减少创建交易的

数量。众包任务的参与者向区块链提交数据时，需要发起一笔交易。如果参与者众多，那么发起的交易费积少成多将成为一笔不小的花费。NF-Crowd 引入了一个上传者角色，所有参与者将数据递交给上传者。上传者将收到的数据组织成一个 Merkle 树，Merkle 树可以保证所有参与者上传的数据均未被修改。最后上传者将 Merkle 树中长度为 32 字节的 Merkle Root 值上传到区块链上。该策略将原本 $O(N)$ 的复杂度降低到了 $O(1)$。

$1\times N$ 型任务优化策略的消耗主要来源于高强度的链上计算以及大量的链上存储。通过计算离链，存储离链可以很好地降低 $1\times N$ 型任务的消耗。

实验结果表明，无论众包项目的规模如何，通过 NF-Crowd 提出的方法，可以将项目的成本降低到 2 美元，这为去中心化的众包解决方案提供了显著的成本效益。

结束语　随着区块链技术不断发展，从最初的区块链 1.0 的比特币，到区块链 2.0 的以太坊，区块链技术的应用也从去中心化的数字加密货币系统发展到现在的去中心化泛智能合约平台。对智能合约的支持使得区块链技术的应用场景得到了大量的扩展。众包作为一种分布式的问题解决机制，通过整合计算机和互联网上具有相应技能的群众来完成计算机单独难以完成的任务。

本文详细综述了区块链＋众包的相关工作，并从智能合约和区块链架构两个角度对相关工作进行了详细的分类讨论，随后总结了区块链众包系统中面临的问题。但是，除了上文所述的技术方面的问题，区块链众包系统的实用性也存在着一些问题，如编写和使用智能合约是需要一定门槛的，区块链众包系统应该为用户扫平这种门槛。

联盟链是区块链技术的一个重要应用。虽然目前利用联盟链平台构建区块链众包系统的相关工作并不多，但是联盟链具有许多公链不具备的特性。如监管机制和通道机制，从隐私保护的角度来说，联盟链比公链更具优势。同时联盟链可拔插的共识机制也使得系统效率和吞吐量得到了提高。但是联盟链的用户基数比不上公链平台（以太坊），可能在项目的推广上会受到较大的阻力，相关的工作也更倾向于在以太坊上进行部署和实验。

去中心化自治社会被认为是区块链技术的下一个发展形态。而基于去中心化的众包系统将社会中的人群和区块链技术相结合，可以认为是去中心化自治社会的初级阶段应用。希望本文工作可以对未来的研究提供参考和启发。

参 考 文 献

[1] NAKAMOTO S. Bitcoin：A peer-to-peer electronic cash system [OL]. http://www.bitcoin.org/bitcoin.pdf，2008.

[2] WOOD G. Ethereum：A secure decentralisedgeneralised transaction ledger[OL]. Ethereum project yellow paper 151 (2014)：1-32. https://www.gavwood.com/paper.pdf.

[3] HOWE J. The rise of crowdsourcing[J]. Wired Magazine，2006，14(6)：1-4.

[4] Upwork，Mountain View，CA，USA. Jun. 2019[OL]. https://www.upwork.com/.

[5] Amazon Mechanical Turk. Jun. 2019[OL]. https://www.mturk.com/.

[6] DIMITRIOS G K，HELEN C L，MICHAEL X，et al. Toward a Blockchain-Enabled Crowdsourcing Platform[J]. IT Professional Magazine，2019，21(5)：18-25.

[7] MAURO C，SANDEEP K E，CHHAGAN L，et al. A Survey on Security and Privacy Issues of Bitcoin[J]. IEEE Communications Surveys and Tutorials，2018，20(4)：3416-3452.

[8] XU X W，WEBER I，STAPLES M，et al. A Taxonomy of Blockchain-Based Systems for Architecture Design [C]// IEEE International Conference on Software Architecture (ICSA'17). IEEE，2017. 2017：243-252.

[9] YANCHAO S. Query Processing on Blockchain Systems [D]. Shanghai：East China Normal University，2020.

[10] ANDROULAKI E，BARGER A，BORTNIKOV V，et al. Hyperleger fabric：a distributed operating system for permissioned blockchains[C]// Proceeding of the 13th EuroSys Conference. 2018，30：1-30.

[11] LI W，FENG C，ZHANG L，et al. A Scalable Multi-Layer PBFT Consensus for Blockchain[J]. IEEE Transactions on Parallel and Distributed Systems，2021，32(5)：1146-1160.

[12] CHASE J M. Quorum whitepaper[R]. White Paper，2016.

[13] SZABO N. Smart Contracts：Building Blocks for DigitalMarkets [OL]. http://www.fon.hum.uva.nl/rob/Courses/InfomationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart_contracts_2.html.

[14] STARK J. Making Sense of Blockchain Smart Contracts [OL]. https://www.coindesk.com/making-sense-smart-contracts/，2016.

[15] WANG S，OUYANG L W，YUAN Y，et al. Blockchain-Enabled Smart Contracts：Architecture，Applications，and Future Trends [J]. IEEE Transactions on Systems Man Cybernetics-Systems，2019，49(11)：2266-2277.

[16] XU X W，PAUTASSO C，ZHUL M，et al. The Blockchain as a Software Connector[C]// 2016 13th Working IEEE/IFIP Conference on Software Architecture (WICSA). 2016：182-191.

[17] FENG J H，LI G L，FENG J H. A Survey on Crowdsoucing[J]. Chinese Journal of Computer，2015，38(9)：1713-1726.

[18] SNOW R，CONNOR B O，JURAFSKY D，et al. Cheap and fast-but is it good? Evaluating non-expert annotations for natural language taks[C]// EMNLP. 2008：254-263.

[19] SEWRJUGIN A，AST F. The crowdjury，a crowdsourced justice system for the collaborationer[OL]. https://medium.com/the-crowdjury/the-crowdjury-a-croedsoured-court-system-for-the-collaboration-ear-66da002750d8. 2015，Accessed June 2019.

[20] YANG P L，LI Q Y，YAN Y B，et al. "Friend is Treasure"：Exploring and Exploiting Mobile Social Contacts for Efficient Task Offloading[J]. IEEE Transactions on Vehicular Technology，2016，65(7)：5485-5496.

[21] JACYNYCZ V，CALVO A，HASSAN S，et al. Betfunding：A distributed bounty-based crowdfunding platformoverethereum [C]// International Conference on Distributed Computing and

Artifical Intelligence(DCAI). 2016:403-411.
[22] ZHU H, ZHOU Z Z. Analysis and outlook of applications of blockchain technology to equity crowdfunding in China[J]. Financial Innovation, 2016, 2(1):29.
[23] AMBATI V, VOGEL S, CARBONELL J. Towards task recommendation in micro-task markets[C]// Proceedings of the 25th AAAI Workshop in Human Computaion. San Francisco, USA, 2011:80-83.
[24] LIU X, LU M, OOI B, et al. CDAS: A crowdsourcing data analytics system[J]. Proceedings of the VLDB Endowment, 2012, 5(10):1040-1051.
[25] YAN Y, ROSALES R, FUNG G, et al. Active Learning from crowds[C]// Proceedings of the 28th International Conference on Machine Learning. Bellevue, USA, 2011:1161-1168.
[26] IPEIROTIS P G, PROVOST F, WANG J. Quality management on amazon mechanical turk[C]// Proceedings of the ACM SIGKDD Workshop on Human Computation. Washington, USA, 2010:64-67.
[27] DAWID A P, SKENE A M. Maximum likelihood estimation of error-rates using the EM algorithm[J]. Applied Statistics, 1979, 28(1):20-28.
[28] Tangle IOTA 2018. Meet the Tangle[OL]. https://www.iota.org/research/meet-the-tangle.
[29] CHURYUMOV A. Byteball: A Decentralized System for Storage and Transfer of Value[OL]. https://byteball.org/Byteball.pdf. 2016.
[30] DEMERS A, GREENE D, HOUSER C, et al. Epidemic algorithms for replicated database maintenance[J]. ACM SIGOPS Operating Systems Review, 1988, 22(1):8-32.
[31] MAYMOUNKOV P, MAZIERES D. Kademlia: A peer-to-peer information system based on the xor metric[C]// International Workshop on Peer-to-Peer Systems. Berlin: Springer, 2002:53-65.
[32] CASTRO M, LISKOV B. Practical Byzantine fault tolerance[C]// OSDI. ACM, New Orleans, USA, 1999:173-186.
[33] LI M, WENG J, YANG A J, et al. CrowdBC: A Blockchain-Based Decentralized Framework for Crowdsourcing[J]. IEEE Transactions on Parallel and Distributed Systems, 2019, 30(6):1251-1266.
[34] TAN L, XIAO H, SHANG X, et al. A Blockchain-based Trusted Service Mechanism for Crowdsourcing System[C]// 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring). 2020:1-6.
[35] GHAFFARIPOUR S, MIRI A. A Decentralized, Privacy-preserving and Crowdsourcing-based Approach to Medical Research[C]// 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC). 2020:4510-4515.
[36] LIN C, HE D B, ZEADALLY S, et al. SecBCS: a secure and privacy-preserving blockchain-based crowdsourcing system [J]. Science China-Information Sciences, 2020, 63(3):20-33.
[37] KADADHA M, MIZOUNI R, SINGH S, et al. ABCrowd: An Auction Mechanism on Blockchain for Spatial Crowdsourcing [J]. IEEE Access, 2020, 8:12745-12757.
[38] GU Y G, CHEN J S, WU X H, et al. An Implement of Smart Contract Based Decentralized Online Crowdsourcing Mechanism [C]// Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence. 2018:195-199.
[39] DING Y, CHEN Z, LIN F, et al. Blockchain-based Credit and Arbitration Mechanisms in Crowdsourcing[C]// 2019 3rd International Symposium on Autonomous Systems (ISAS). 2019:490-495.
[40] WANG S, TAHA A F, WANG J. Blockchain-Assisted Crowdsourced Energy Systems[C]// 2018 IEEE Power & Energy Society General Meeting (PESGM). 2018:1-5.
[41] LIU K, CHEN W, ZHANG Z. Blockchain-Empowered Decentralized Framework for Secure and Efficient Software Crowdsourcing[C]// 2020 IEEE World Congress on Services (SERVICES). Beijing, China, 2020.
[42] LU Y, TANG Q, WANG G. On Enabling Machine Learning Tasks atop Public Blockchains: A Crowdsourcing Approach [C]// 2018 IEEE International Conference on Data Mining Workshops (ICDMW). 2018:81-88.
[43] WU Y, TANG S, ZHAO B, et al. BPTM: Blockchain-Based Privacy-Preserving Task Matching in Crowdsourcing[J]. IEEE Access, 2019, 7:45605-45617.
[44] SHENG D, XIAO M, LIU A, et al. CPchain: A Copyright-Preserving Crowdsourcing Data Trading Framework Based on Blockchain[C]// 2020 29th International Conference on Computer Communications and Networks. 2020.
[45] LIN Y, ZHANG C. Crowdsourcing System Based on Blockchain [C]// 2020 IEEE 11th International Conference on Software Engineering and Service Science (ICSESS). 2020:98-101.
[46] CHEN Y, YIN H, XIANG Y, et al. CVT: A Crowdsourcing Video Transcoding Scheme Based on Blockchain Smart Contracts [J]. IEEE Access, 2020, 8:220672-220681.
[47] HAN S Y, XU Z H, ZENG Y X, et al. FLUID: A Blockchain based Framework for Crowdsourcing[C]// Proceedings of the 2019 International Conference on Management of Data, New York: Assoc Computing Machinery. 2019:1921-1924.
[48] ZHANG W K, HONG Z C, CHEN W H. Hierarchical Pricing Mechanism With Financial Stability for Decentralized Crowdsourcing: A Smart Contract Approach[J]. IEEE Internet of Things Journal, 2021, 8(2):750-765.
[49] GAO L P, CHENG T, GAO L. TSWCrowd: A Decentralized Task-Select-Worker Framework on Blockchain for Spatial Crowdsourcing[J]. IEEE Access, 2020, 8:220682-220691.
[50] LU Y, TANG Q, WANG G L, et al. ZebraLancer: Private and Anonymous Crowdsourcing System atop Open Blockchain[C]// IEEE 38th International Conference on Distributed Computing Systems. New York: IEEE, 2018:853-865.
[51] FENG W, YAN Z. MCS-Chain: Decentralized and trustworthy mobile crowdsourcing based on blockchain[J]. Future Generation Computer Systems-the International Journal of Escience, 2019, 95:649-666.
[52] XU X L, LIU Q X, ZHANGX Y, et al. A Blockchain-Powered Crowdsourcing Method With Privacy Preservation in Mobile

Environment[J]. IEEE Transactions on Computational Social Systems，2019，6(6)：1407-1419.

[53] ZHU X Y，LI Y，FANG L，et al. An Improved Proof-of-Trust Consensus Algorithm for Credible Crowdsourcing Blockchain Services[J]. IEEE Access，2020，8：102177-102187.

[54] YU Y，LIU S M，GUO L，et al. CrowdR-FBC：A Distributed Fog-Blockchains for Mobile Crowdsourcing Reputation Management[J]. Ieee Internet of Things Journal，2020，7(9)：8722-8735.

[55] FERNANDEZ-CARAMES T M，FROIZ-MIGUEZ I，BLANCO-NOVOA O，et al. Enabling the Internet of Mobile Crowdsourcing Health Things：A Mobile Fog Computing，Blockchain and IoT Based Continuous Glucose Monitoring System for Diabetes Mellitus Research and Care[J]. Sensors，2019，19(15)：24.

[56] ZHU S，HU H，LI Y，et al. Hybrid Blockchain Design for Privacy Preserving Crowdsourcing Platform[C]// 2019 IEEE International Conference on Blockchain (Blockchain). 2019：26-33.

[57] ZHAO Y，ZHAO J，JIANG L，et al. Privacy-Preserving Blockchain-Based Federated Learning for IoT Devices[J]. IEEE Internet of Things Journal，2021，8(3)：1817-1829.

[58] YANG H，WANG G，ZHAI Z，et al. Towards Decentralized Trust Management Using Blockchain in Crowdsourcing Networks[C]// 2020 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC). 2020：23-28.

[59] ZHU S，CAI Z，HU H，et al. zkCrowd：A Hybrid Blockchain-Based Crowdsourcing Platform[J]. IEEE Transactions on Industrial Informatics，2020，16(6)：4196-4205.

[60] ZHU X Y. Formal analysis of the DAO exploit[J]. Information Technology And Network Security，2021，40(5)：13-19.

[61] HAN X，YUAN Y，WANG F Y. Security Problems on Blockchain：The State of the Art and Future Trends[J]. Acta Automatica Sinica，2019，45(1)：208-225.

[62] LUU L. Making Smart Contracts Smarter[C]// Paper presented at Computer and Communications Security. 2016：254-269.

[63] TORRES C F，SCHUTTE J，STATE R，et al. OSIRIS：Hunting for Integer Bugs in Ethereum Smart Contracts[C]// ACSAC. New York：Assoc Computing Machinery，2018：664-676.

[64] TSANKOV P，DAN A，DRACHSLER-COHEN D，et al. Securify：Practical Security Analysis of Smart Contracts[C]// The 2018 ACM SIGSAC Conference. ACM，2018.

[65] GRISHCHENKO I，MAFFEI M，SCHNEIDEWIND C. A semantic framework for the security analysis of ethereum smart contarcts[C]// Proceedings of 2018 International Conference on Principles of Security and Trust. Thessaloniki，Greece：Springer，2018：243-269.

[66] ZHAO H，LI X，TAN J C，et al. Research status of smart contract security[J]. Information Technology And Network Security，2021，40(5)：1-6，19.

[67] AU M H. A New Payment System for Enhancing Location Privacy of Electric Vehicles[J]. IEEE Transactions on Vehicular Technology，2014，63(1)：3-18.

[68] NOETHER S. Ring signature confidential transactions for Monero [OL]. https://eprint.iacr.org/2015/ 1098.

[69] KYUNG-AH S. An efficient ring signature scheme from parings [J]. Infomation Sciences，2015，300：63-69.

[70] MIERS I，GARMAN C，GREEN M，et al. Zerocoin：anonymous distributed e-cash from bitcoin[C]// 2013 IEEE Symposium on Security and Privacy. 2013：397-411.

[71] BITANSKY N，CHIESA A，ISHAI Y，et al. Succinct non-interactive arguments via linear interactive proofs[C]// Proceedings of the 2013 Theory of Cryptography. 2013：315-333.

[72] SASSON E B，CHIESA A，GARMAN C，et al. Zerocash：Decentralized Anonymous Payments from Bitcoin[C]// 2014 IEEE Symposium on Security and Privacy. 2014：459-474.

[73] ZHANG H R，WANG M Q，LI G S. The development status of frontier technology of blockchain security and privacy protection [J]. Information Technology And Network Security，2021，40(5)：7-12.

[74] LIU F，YANG J，LI Z B，et al. A secure Multi-Party Computation Protocol for Universal Data Privacy Protection Based on Blockchain[J]. Journal of Computer Research and Development，2021，58(2)：281-290.

[75] DUAN H，ZHENG Y，DU Y，et al. Aggregating Crowd Wisdom via Blockchain：A Private，Correct，and Robust Realization[C]// 2019 IEEE International Conference on Pervasive Computing and Communications. 2019：1-10.

[76] CHEN T，LI Z，ZHOU H，et al. Towards Saving Money in Using Smart Contracts[C]// 2018 IEEE/ACM 40th International Conference on Software Engineering：New Ideas and Emerging Technologies Results (ICSE-NIER). 2018：81-84.

[77] LI C，PALANISAMY B，XU R，et al. NF-Crowd：Nearly-free Blockchain-based Crowdsourcing[C]// 2020 International Symposium on Reliable Distributed Systems (SRDS). 2020：41-50.

[78] BENET J. Ipfs-content addressed，versioned，p2p file system [J]. arXiv：1407.3561，2014.

LI Yu，born in 1989，Ph.D，post-doctoral researcher，is a member of China Computer Federation. Her main research interests include crowdsourcing，spatial recommendation，database optimization，and cloud computing.

YIN Yu-yu，born in 1980，Ph.D professor，master supervisor，is a member of China Computer Federation. His main research interests include edge computing，service computing and formal method.