面向 6G 的区块链物联网数据共享和存储机制

蒋宇娜[1]，葛晓虎[1]，杨旸[2,3]，王承祥[4,5]，李颉[6]

（1. 华中科技大学电子信息与通信学院，湖北 武汉 430074；2. 上海科技大学信息科学与技术学院，上海 201210；
3. 鹏城实验室网络通信研究中心，广东 深圳 518000；4. 东南大学移动通信国家重点实验室，江苏 南京 210096；
5. 紫金山实验室，江苏 南京 211111；6. 上海交通大学电子信息与电气工程学院，上海 200240）

摘 要：考虑不同物联网系统的异构性以及集中化数据处理平台单点故障等问题，提出一种基于区块链技术的去中心化物联网数据共享和存储方案。通过存储证明（PoS）的共识机制，实现区块共识和共享数据的分布式存储。基于 Gossip 协议提出一种在区块链网络的共识节点和验证节点间的区块分层传播机制。然后，推导了区块传播时延模型以及区块链网络的去中心化评估模型，同时对区块传播时延以及网络去中心化程度进行均衡化分析。仿真结果表明，随着共识节点的能力最小值增加，区块传播时延、区块链网络去中心化程度减小。作为应用示例，针对确诊患者轨迹数据共享场景，基于以太坊开发平台进行了数据共享智能合约的实现和测试。

关键词：区块链；数据共享；数据存储；区块传播时延；去中心化

中图分类号：TN92

文献标识码：A

doi: 10.11959/j.issn.1000−436x.2020211

6G oriented blockchain based Internet of things data sharing and storage mechanism

JIANG Yu’na[1], GE Xiaohu[1], YANG Yang[2,3], WANG Chengxiang[4,5], LI Jie[6]

1. School of Electronic Information and Communication, Huazhong University of Science and Technology, Wuhan 430074, China
2. School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China
3. Research Center for Network Communication, Peng Cheng Laboratory, Shenzhen 518000, China
4. National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China
5. Purple Mountain Laboratories, Nanjing 211111, China
6. School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

Abstract: Considering the heterogeneity of various IoT system and the single point failure of centralized data-processing platform, a decentralized IoT data sharing and storage method based on blockchain technology was proposed. The block consensus and decentralized storage of shared data were realized through the PoS consensus mechanism. A block layered propagation mechanism between consensus node and verified node was proposed based on the Gossip protocol. The block propagation delay model and decentralization evaluation model of blockchain networks were derived. The trade-off between the block propagation delay and the decentralization degree of networks was analyzed. The simulation results demonstrate that the block propagation delay and degree of network decentralization decrease with the increase of minimal capabilities of consensus nodes. As an application example, in the trajectory data sharing scenario of confirmed patients, the data sharing smart contract is implemented and tested based on the Ethereum development platform.

Key words: blockchain, data sharing, data storage, block propagation delay, decentralization

---

收稿日期：2020−08−19；**修回日期**：2020−10−10

通信作者：杨旸，yangyang@shanghaitech.edu.cn

基金项目：国家重点研发计划基金资助项目（No.SQ2020YFB210291）；国家自然科学基金资助项目（No.61932014）；华中科技大学研究生创新基金资助项目（No. 2020yjsCXCY060）

Foundation Items: The National Key Research and Development Program of China (No.SQ2020YFB210291), The National Natural Science Foundation of China (No.61932014), Graduates’ Innovation Fund, Huazhong University of Science and Technology (No.2020yjsCXCY060)

1　引言

6G 移动通信将促进物联网时代的全面实现。6G 移动通信网络不仅连接人，还连接计算资源、车辆、设备、传感器和机器人代理等，以满足完全互联、智能数字世界的需求[1-4]。物联网是一种将现实世界中的物体与网络连接起来的网络范式[5]。物联网允许设备在无人干预的情况下进行数据的收集、处理和通信[6-7]。在 6G 移动通信系统中，物联网的全面部署使网络接入数快速增长。根据爱立信的预测，到 2025 年将会有超过 249 亿台设备连接到网络。智能设备量的增长使网络数据量呈现爆炸式增长态势[8-10]。通过对物联网数据的收集和分析，可以进一步挖掘物联网数据的潜在价值[11]。但是因为不同的物联网系统之间存在数据壁垒，使数据的价值被抑制。例如在新冠肺炎疫情暴发阶段，为了尽可能准确地找到潜在的病毒携带者，很多互联网公司推出了确诊患者的全国同乘查询服务，但是这些服务无法覆盖一些公共场所，如商场、广场、公园等，从而无法获得确诊患者的完整轨迹。在万物互联时代，如何集合不同物联网系统采集的数据，例如公共场所的摄像头、不同的应用程序等，实现多个物联网系统之间的数据共享，从而确定确诊患者的完整轨迹是一个难题。

如果将所有的数据都发送到集中式云平台进行处理，会带来巨大的挑战。首先，如果中央服务器出现故障，整个网络服务器都会面临瘫痪风险，例如对集中式服务器进行拒绝服务攻击可能导致单点故障问题。其次，用户对个人数据如何使用以及被谁使用的控制是有限的，存储在集中式服务器中的数据可能会泄露个人隐私。最后，存储在集中式云中的数据缺乏可靠性和可跟踪性。集中式物联网基础设施要求信任第三方进行数据处理，而存储在集中式服务器上的数据有被删除或篡改的风险[12]。区块链技术因分散自治、不可篡改、可溯源等特点在近些年被广泛关注。区块链技术被认为是简化网络管理并提升 6G 网络性能的关键去中心化技术[13]。区块链上存储的数据需由全网共同维护，可以在缺乏信任的节点之间有效地传递价值[14]。利用区块链技术，以前只能通过可信第三方平台进行的物联网数据共享现在可以通过去中心化的方式运行[15]。但是，基于区块链技术的物联网数据共享研究仍然面临很多挑战，其中一个关键问题是共享数据的存储问题。

目前，很多基于区块链的数据共享研究都忽略了共享数据的存储问题。文献[16]为了解决无信任环境中医疗数据的共享问题，使用区块链技术为大数据实体间的共享医疗数据提供数据来源、审计和控制。文献[17]针对多云平台中的数据安全共享问题，提出一个基于区块链和智能合约的可靠协作模型，同时分析了参与者之间的拓扑关系，并在收益分配过程中建立了由简单到复杂的 Shapley 值模型。文献[18]提出了一种基于区块链的高效数据收集和安全共享方案，结合以太坊区块链和深度强化学习创建一个可靠和安全的环境，其中深度强化学习实现收集数据量的最大化，区块链技术用来保证数据共享的可靠性和安全性。文献[19]将区块链和支持向量机结合，提出一种可以保护物联网数据隐私的数据训练方案。通过区块链技术，可以在多个数据提供商之间构建一个安全可靠的数据共享平台。在文献[16-19]提出的数据共享方案中，数据拥有者将需要共享的数据加密后发送到区块链上进行存储。区块链分布式账本记录在每个全节点中，如果将共享的数据存储在区块链上将会造成存储资源的极大浪费。另外，如果交易包含的数据量过大，对于整个区块链网络的性能也会产生消极影响。除此之外，区块链的透明性很容易泄露共享数据的隐私性。所以，在基于区块链的物联网数据共享中，将共享的数据存储在区块链上并不是一种可行的方案。另外，还有一些数据共享的研究中将星际文件系统（IPFS, inter planetary file system）用于共享数据的存储。文献[20]研究了分布式存储系统的数据存储与共享方案，提出了将 IPFS、以太坊和基于属性的加密技术相结合的架构。文献[21]基于区块链技术和 IPFS 提出了一个新型电子医疗数据共享框架，并使用智能合约设计了一个可靠的访问控制机制，以实现不同患者和医疗供应商之间安全的电子医疗数据共享。文献[22]为了在 IPFS 存储环境下实现电子病历的安全存储和高效共享，构造了一种基于属性的加密方案。通过将加密的电子医疗数据存储在分布式的 IPFS 中，保证了存储平台的安全性，避免了单点失效的问题。通过 IPFS 将共享的数据转化成哈希值，然后将哈希值存储在区块链上，从而避免了大量共享数据的链上存储。但是 IPFS 网络要想稳定运行，需要用户贡献存储空间和网络带宽，如果没有恰当的奖励机制，那么巨大的

资源开销很难维持网络的持久运行。由此可见，基于区块链技术的数据共享和存储是一个值得研究的课题。

在物联网中，使用区块链技术实现数据的去中心化共享和存储对物联网本身来说是一个巨大的挑战。因为物联网中的大多数设备为低功耗设备，不具备参与区块链网络分布式共识的能力。在区块链的共识机制中，参与共识过程的节点需要负责共识的形成、交易验证以及区块的验证和打包。区块链网络中设备的能力是设计共识机制时需要考虑的主要因素之一。物联网中资源受限的节点，如传感器，无法承担共识任务。而对于设备能力比较强的节点（如网关等），仍然可以在区块链的共识过程中发挥着重要的作用。除此之外，共识节点的比例会对区块链网络的性能产生影响，如果共识节点的比例过小，则整个区块链网络的分散化程度较小，不利于整个系统的安全；如果共识节点的比例过大，则区块传播时延就会很大，在需要大量数据交互的场景下，不能满足服务需求，例如，在新冠肺炎疫情暴发时，需要不同的物联网系统之间共享大量的确诊患者数据，较大的共识时延会直接导致数据共享效率低下。

基于上述问题，本文提出了一种基于区块链技术的物联网数据共享和存储方案，主要贡献如下。

1) 提出一种基于区块链技术的数据共享和存储框架，实现 6G 时代物联网数据的去中心化共享和存储。通过存储证明（PoS, proof of storage）的共识机制，将区块共识和共享数据的分布式存储相结合。

2) 基于 Gossip 协议提出区块的分层传播机制。通过对物联网设备能力的分析，推导了区块传播时延模型和区块链网络的去中心化评估模型。

3) 仿真分析表明，本文提出的基于 Gossip 协议的分层传输方案相较于传统的 Gossip 协议传输在区块传播时延上有大幅降低。区块传播时延、区块链网络去中心化程度随着共识节点的能力最小值增大而减小。当共识节点计算能力阈值为 0.535×10^4 Hz 或 1.015×10^4 Hz、共识节点存储能力阈值为 106 GB 或 315 GB 时，区块链网络去中心化程度与区块传播时延达到均衡。

4) 针对确诊患者轨迹数据共享场景，基于以太坊开发平台进行数据共享智能合约的实现和测试。

2 系统模型

2.1 基于区块链的物联网数据共享和存储框架

基于区块链技术的 6G 时代物联网数据共享和存储系统框架如图 1 所示。本文根据物联网设备的计算能力和存储能力将设备分为共识节点、验证节点和其他节点。共识节点参与交易的广播、验证，区块的打包和共识中，验证节点负责交易的广播和验证。在共识节点和验证节点上都保存了完整的分布式账本。系统框架主要包括智能合约、区块链网络、分布式账本和物联网设备。

1) 智能合约。智能合约是运行在区块链上模块化、可重用、自动执行的脚本。智能合约允许节点以可验证的方式在区块链网络上执行脚本。由于智能合约保存在区块链上，因此每个节点都可以查看和执行其指令，并查看与智能合约交互的日志[23]。通过智能合约，物联网中的数据拥有者可以将数据存储在去中心化的网络中，数据需求者和数据拥有者可以自动进行数据共享交易。

2) 区块链网络。区块链网络中主要包括共识节点和验证节点。共识节点在去中心化网络中具有重要作用，除了需要验证数据共享交易或者数据存储交易，执行共识算法，还需要为数据拥有者提供存储空间以存储共享数据。验证节点需要验证交易和区块，帮助共识节点执行共识过程。

3) 分布式账本。区块链本质上是一个不可篡改的去中心化数字账本，并且由区块链网络中的共识节点和验证节点共同维护。区块链分布式账本上记录了物联网不同节点间的交互日志。

4) 物联网设备。物联网设备主要包括数据拥有者进行数据收集的设备。6G 时代将会实现空-天-地一体化，从而会有更多的物联网设备接入网络，包括无人机、多样化的智能设备、自动驾驶汽车和深海潜水艇等。数据拥有者可以通过多样化的物联网设备进行数据的采集。当数据需求者在区块链上提交数据需求后，数据拥有者根据数据需求方的数据需求将采集到的数据提供给数据需求者。

物联网设备间数据的共享和存储流程如下：当数据需求者需要从其他的设备获得数据时，数据需求者会通过智能合约在区块链上发布数据请求的交易。数据拥有者在区块链网络上监听到数据需求者的交易后，如果有符合要求的数据，则会执行下

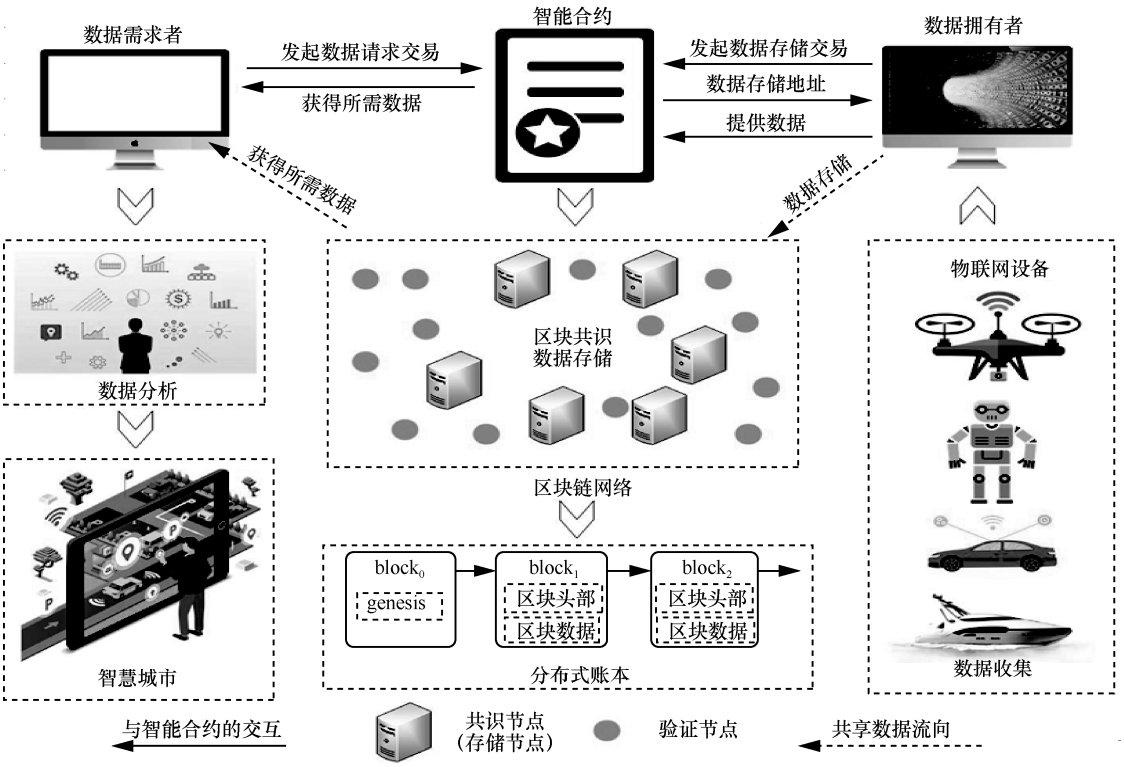


图 1　系统框架

述操作：1) 通过智能合约向共识节点租用存储资源；2) 得到提供存储资源的共识节点的响应后，将加密后的数据、对数据的描述、存储的时间和需要存储的副本数发送给共识节点，共识节点将共享数据的地址以及数据的描述发送到区块链上。通过共享数据存储地址链上存储、数据链下存储的方式，数据拥有者可以根据实际需求决定共享数据存储在共识节点处的时间。数据需求者从数据拥有者那里获取数据密钥后，就可以从相应的数据地址获取需求的数据。数据需求者获取数据后，对数据进行分析使用，从而更好地挖掘数据的潜在价值。

2.2　共识机制

区块链中一个核心概念就是去中心化，区块链网络中没有和传统数据库系统一样的中心数据库。每个节点都是对等的，从而需要共识机制保证所有对等节点之间可以有效地协作[24]。共识机制是区块链事务达成分布式共识的算法。比特币和以太坊都采用对计算强依赖的工作量证明（PoW, proof of work）算法。物联网中存在大量的低功耗设备，计算资源的匮乏使高难度的 PoW 算法不再适用于物联网数据共享场景。Hyperledger Fabric 采用传统的拜占庭容错算法，如实用拜占庭容错（PBFT, practical Byzantine fault tolerance）算法。在存有大量节点的物联网场景中，PBFT 的通信复杂度将会大幅增加[25]。

考虑到数据共享中的存储需求，本文采用 PoS 共识机制[26]。共识节点成功在区块链上添加区块的概率是由当前共识周期内的存储空间占全网存储空间的比值决定的。令 N_c 为网络中共识节点的数量，p_i^t 为第 t 个共识周期内共识节点 i 为网络提供的时空证明容量，则共识节点 i 在 t 个共识周期内为网络提供的有效存储空间占比，即存储算力 pos_i^t 为

$$\mathrm{pos}_i^t = \frac{p_i^t}{\sum_{j=1}^{N_c} p_j^t} \tag{1}$$

存储算力 pos_i^t 越大，则共识节点成功在区块链上添加区块的概率越大。与 PoW 共识机制相比，PoS 共识机制使共识节点之间的共识过程不需要浪费大量的算力完成无意义的哈希计算任务。同时，PoS 共识机制也可以激励共识节点进行共享数据的存储，因为在一个共识周期内，存储数据量越多，共识节点获得共识奖励的概率就越大。为了增加区块链网络的去中心化程度，区块的共识结果不仅由共识节点决定，还需要由验证节点决定。当共识节

点生成区块后，需要将产生的区块传播给所有的共识节点和部分验证节点进行验证，经过验证后的区块可以被添加到区块链上。

2.3 区块传播机制

区块的传播机制如图 2 所示。当共识节点 i 产生一个新的区块 block_i 时，共识节点 i 需要将区块 block_i 传输给其他的共识节点进行验证，同时还需要招募一部分的验证节点进行验证。接收到这个区块的其他共识节点也需要分别招募验证节点进行验证。本文的区块传播采用基于 Gossip 协议的分层传播，包括共识节点层和验证节点层，即共识节点之间、共识节点和验证节点之间以及验证节点之间都采用 Gossip 协议进行区块传播。Gossip 协议最早在文献[27]中被提出，主要用在分布式数据库系统中各个副本节点之间的数据同步，其基本思想是节点随机地选择一些节点进行信息传递，接收到信息的节点会采用同样的方式把信息传递给其他节点。节点间区块的传播都采用无线多播的方式。区块 block_i 传播过程如下：1) 共识节点之间：共识节点 i 将区块 block_i 传输给其他共识节点；2) 共识节点和验证节点之间：共识节点 i 将区块 block_i 传输给其所招募的验证节点；3) 验证节点之间：区块 block_i 在共识节点 i 招募的验证节点间进行传输。令共识节点集合为 $C=\{c_1,c_2,\cdots,c_{|C|}\}$ 且 $|C|=N_c$，验证节点的集合为 $V=\{v_1,v_2,\cdots,v_{|V|}\}$ 且 $|V|=N_v$，其中 $N_v>N_c$。假设每个共识节点招募的验证节点的个数相等，且等于 αN_v $(0<\alpha<1)$，则区块 block_i 需要被验证的总次数为 $N_c\left(\alpha N_v+1\right)-1$。

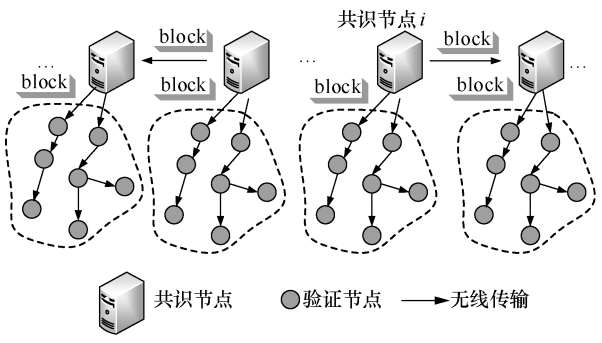


图 2 区块的传播机制

3 区块传播时延

本节结合物联网节点的计算能力和存储能力，分析区块链网络中的区块传播时延。

在基于区块链的物联网数据共享和存储中，区块链的分布式共识由共识节点和验证节点共同完成。本文使用帕累托分布[28]描述物联网设备的计算能力和存储能力。帕累托分布最初用来描述社会的财富状况，现在帕累托分布被拓展到更广泛的范围，可以用来描述使用传输控制协议的网络流量的文件大小分布（即多数较小的文件以及少数较大的文件）、人类居住区的大小（即少数的城市以及多数的小村庄）等。考虑到物联网中存在大量的低功耗设备以及少量能力较强的设备，本文使用帕累托分布来描述物联网设备的能力。令物联网设备计算能力为 $\{X_1,X_2,\cdots,X_N\}$，且 X_i 服从参数为 ζ 和 σ_c 的帕累托分布，其中 $\sigma_c=\min\{X_i\}$；令物联网设备的存储能力为 $\{Y_1,Y_2,\cdots,Y_N\}$，且 Y_i 服从参数为 ζ 和 σ_s 的帕累托分布，其中 $\sigma_s=\min\{Y_i\}$；N 为物联网设备的总数量（$N\gg N_c+N_v$），则物联网设备计算能力 X 的生存函数为

$$\overline{F_c}(x)=\Pr(X>x)=\left[\frac{x}{\sigma_c}\right]^{-\zeta},\ x\geqslant\sigma_c>\zeta \tag{2}$$

物联网设备计算能力 X 的概率密度函数为

$$f_X(x)=\begin{cases}\dfrac{\zeta\sigma_c^{\zeta}}{x^{\zeta+1}}, x\geqslant\sigma_c\\ 0,\ x<\sigma_c\end{cases} \tag{3}$$

物联网设备存储能力 Y 的生存函数为

$$\overline{F_s}(y)=\Pr(Y>y)=\left[\frac{y}{\sigma_s}\right]^{-\zeta},\ y\geqslant\sigma_s>\zeta \tag{4}$$

令共识节点计算能力阈值为 X_c，共识节点存储能力阈值为 Y_c。当节点计算能力大于 X_c 且存储能力大于 Y_c 时，物联网节点可作为共识节点，则物联网节点为共识节点的概率为

$$F_{\text{con}}=\overline{F_c}(X_c)\overline{F_s}(Y_c)=\left[\frac{X_c}{\sigma_c}\right]^{-\zeta}\left[\frac{Y_c}{\sigma_s}\right]^{-\zeta}$$
$$X_c\geqslant\sigma_c>\zeta\text{且}Y_c\geqslant\sigma_s>\zeta \tag{5}$$

其中，参数 ζ 越大，物联网设备中共识节点的比例就越小。共识节点的数量为

$$N_c=NF_{\text{con}} \tag{6}$$

共识节点计算能力的期望为

$$\mathrm{E}(X)_{\text{con}}=\int_{X_c}^{+\infty}\frac{\zeta X_c^{\zeta}}{x^{\zeta}}\mathrm{d}x \tag{7}$$

令验证节点计算能力阈值为 X_v，验证节点存储能力阈值为 Y_v。当节点的计算能力大于阈值 X_v 且小于阈值 X_c、存储能力大于阈值 Y_v 且小于阈值 Y_c 时，物联网节点可作为验证节点，则物联网节点为验证节点的概率为

$$\begin{aligned}&F_{\text{ver}}=\Pr\left(X_v<X<X_c\right)\Pr\left(Y_v<Y<Y_c\right)=\\&\left[\Pr\left(X>X_v\right)-\Pr\left(X>X_c\right)\right]\cdot\\&\left[\Pr\left(Y>Y_v\right)-\Pr\left(Y>Y_c\right)\right]=\\&\left[\left(\frac{X_v}{\sigma_c}\right)^{-\zeta}-\left(\frac{X_c}{\sigma_c}\right)^{-\zeta}\right]\left[\left(\frac{Y_v}{\sigma_s}\right)^{-\zeta}-\left(\frac{Y_c}{\sigma_s}\right)^{-\zeta}\right]\\&X_c>X_v\geqslant\sigma_c>\zeta\text{且}Y_c>Y_v\geqslant\sigma_s>\zeta\end{aligned}\tag{8}$$

验证节点计算能力的期望为

$$\mathrm{E}\left(X\right)_{\text{ver}}=\int_{X_v}^{X_c}\frac{\zeta X_v^{\zeta}}{x^{\zeta}}\mathrm{d}x\tag{9}$$

$X=\beta_1 f_m$ 为物联网节点的计算能力，其中 f_m 是节点的 CPU 频率。$Y=\beta_2 e$ 为物联网节点的存储能力，其中 e 是节点的内存大小。验证节点的数量为

$$N_v=NF_{\text{ver}}\tag{10}$$

在区块传播过程中，节点 A 和节点 B 之间区块的传输协议采用传统块传播协议[29]，实现过程如图 3 所示。在发送区块之前，节点 A 发送一个 Inventory 的信息给节点 B，验证节点 B 是否已经拥有区块，如果节点 B 没有该区块，则区块 B 会回复给节点 A 一个 Getdata 的信息，并等待接收区块 block 的信息。通过块传播协议，节点可以减少不必要的信息传输[30]。

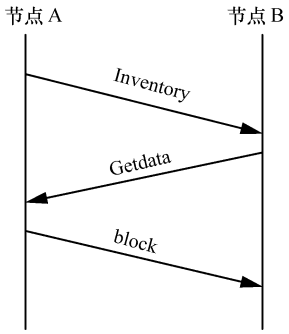


图 3　区块在节点 A 和节点 B 之间的传输

区块的传播时延包括 3 个部分：区块的传输时延、区块的验证时延以及节点之间交换 Inventory 和 Getdata 信息的时延。交换 Inventory 和 Getdata 信息的平均往返时间表示为 τ_{RTT} [31]。节点之间的区块传输采用 Gossip 协议。对于一个 N 个节点的网络，假设在每个 Gossip 周期，接收到区块的节点都能至少再将区块传输给一个节点，那么区块传输到 N 个节点的周期数为 $\log(N)$。当区块大小为 s 时，区块的传输时延可以表示为[32]

$$\tau_{p,b}=\frac{s}{c}\log\left(N\right)\tag{11}$$

其中，c 是各链路的平均有效信道容量。对于区块，单个节点的验证时延为

$$\tau_{1,v,b}=\frac{sU}{f_m}\tag{12}$$

其中，U 表示每比特信息所需要的 CPU 周期数。令共识节点的 CPU 频率为 $f_{m,c}$，验证节点的 CPU 频率为 $f_{m,v}$，则区块在全网中传播时的总验证时延为

$$\begin{aligned}&\tau_{v,b}=\alpha N_v\frac{sU}{f_{m,c}}+\left(N_c-1\right)\frac{sU}{f_{m,v}}=\\&\alpha N_v\frac{\beta_1 sU}{\mathrm{E}\left(X\right)_{\text{ver}}}+\left(N_c-1\right)\frac{\beta_1 sU}{\mathrm{E}\left(X\right)_{\text{con}}}\end{aligned}\tag{13}$$

因此，区块 block 的平均传播时延可以表示为

$$\begin{aligned}&\tau=\tau_{p,b}+\tau_{v,b}+2\left(N_c+\alpha N_v-1\right)\tau_{\text{RTT}}=\\&\frac{s}{c}\left[\log\left(N_c+\alpha N_v-1\right)\right]+\alpha N_v\frac{\beta_1 sU}{\mathrm{E}\left(X\right)_{\text{ver}}}+\\&\left(N_c-1\right)\frac{\beta_1 sU}{\mathrm{E}\left(X\right)_{\text{con}}}+2\left(N_c+\alpha N_v-1\right)\tau_{\text{RTT}}\end{aligned}\tag{14}$$

4　区块链网络去中心化程度

去中心化是区块链网络的重要特性之一，即在理想情况下任何一个或者少数实体都不具备操控、颠覆整个区块链网络的权力。现有研究中提出将中本系数作为一种衡量区块链网络去中心化的指标。基于经济学中衡量国民财富分布的基尼系数和洛伦兹曲线，本文首先把区块链系统拆分成多个关键的子系统，然后分别计算出能够操纵每个子系统的最少实体数量，最后把需要实体数量的最小值作为整个系统的去中心化指数。数值越高的系统去中心化程度越高。文献[33]使用信息熵的方法衡量区块链系统的分散化程度，并分别计算了比特币和以太坊系统中挖掘到的区块以及账户余额的熵。计算结果表明，比特币的分散化程度高于以太坊的分散化程度。

在本文基于区块链技术的物联网数据共享和存储中，参与共识过程和验证过程的节点对于区块链网络的去中心化影响要强于其他未参与共识过程和验证过程的节点。共识节点负责将交易进行打包、广播并形成共识，同时需要存储共享的数据，在区块链网络的共识过程担任重要的角色。验证节点需要对共识节点打包的交易进行验证，只有通过一定比例节点验证的区块才会被添加到区块链上。因此，验证节点同样影响着区块链网络的去中心化程度。共识节点存储数据量的差异性也影响着区块链网络的去中心化程度。存储数据量的差异性越大，区块链网络越趋向于集中化，因此，本文主要从共识节点之间数据存储的差异性、共识节点的比例和验证节点的比例这 3 个方面进行区块链网络的去中心化建模分析。

共识节点之间数据存储的差异性主要需要考虑共享的数据文件如何在共识节点之间进行存储。假设在一个共识周期内，需要共享的数据文件为 $F=\{f_1,f_2,\cdots,f_{|F|}\}$，且 $|F|=N_F$。本文假设数据文件的大小相等，使用参数 $x_{c,f}\in\{0,1\}$ 表示数据文件 $f\in F$ 是否会存储在共识节点 $c\in C$ 上，$x_{c,f}=0$ 表示数据文件 f 没有存储在共识节点 c 上，$x_{c,f}=1$ 表示数据文件 f 存储在共识节点 c。在一个共识周期内，共享数据文件在共识节点上的存储情况表示为

$$\boldsymbol{\Omega}=\begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,|F|} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,|F|} \\ \vdots & \vdots & \ddots & \vdots \\ x_{|C|,1} & x_{|C|,2} & \cdots & x_{|C|,|F|} \end{bmatrix} \tag{15}$$

则存储在共识节点 c_i 处的数据文件份数为

$$\Omega_{i,|F|}=x_{i,1}+x_{i,2}+\cdots+x_{i,|F|} \tag{16}$$

为了防止存储的数据文件丢失，数据文件 f_j 可以在多个共识节点上进行存储，则数据文件 f_j 存储的副本数量为

$$\Omega_{|C|,j}=x_{1,j}+x_{2,j}+\cdots+x_{|C|,j} \tag{17}$$

共识节点之间存储数据量的差异性使用共识节点之间存储数据文件数的方差 var 表示。var 越大，表明共识节点之间存储的共享数据文件数差异越大。var 表示为

$$\mathrm{var}=\frac{\sum\left(\Omega_{i,|F|}-\dfrac{\sum\sum x_{i,j}}{N_c}\right)^2}{N_c-1} \tag{18}$$

区块链网络的去中心化程度 D 为

$$D=\frac{\beta_3 F_{\mathrm{con}}+\beta_4 F_{\mathrm{ver}}}{\mathrm{var}}+\kappa_2 \tag{19}$$

其中，β_3 和 β_4 为系统参数，κ_2 为误差参数。D 越大，区块链网络的去中心化程度越大。

为了对去中心化程度 D 与区块传播时延 τ 进行均衡化分析，需要对区块传播时延 τ 和去中心化程度 D 进行 min-max 归一化处理。区块传播时延 τ 进行归一化处理后为

$$\tau_{\to 1}=\frac{\tau-\tau_{\mathrm{mean}}}{\tau_{\max}-\tau_{\min}} \tag{20}$$

其中，τ_{mean} 是区块传播平均时延，$\tau_{\max}$ 是区块传播时延最大值，$\tau_{\min}$ 是区块传播时延最小值。去中心化程度 D 进行 min-max 归一化处理后为

$$D_{\to 1}=\frac{D-D_{\mathrm{mean}}}{D_{\max}-D_{\min}} \tag{21}$$

其中，D_{mean} 是 D 的平均值，$D_{\max}$ 是 D 的最大值，$D_{\min}$ 是 D 的最小值。令 $\phi=\tau_{\to 1}-D_{\to 1}$，可以根据系统的具体性能需求对参数 ϕ 的数值进行调整。对于时延要求较高的场景，取 $\phi>0$；对于去中心化程度要求较高的场景，取 $\phi<0$。

5 实验结果

本节首先仿真分析了共识节点的能力与区块链网络中的区块传播时延以及网络去中心化程度的关系。然后，针对确诊患者轨迹数据共享场景，基于以太坊开发平台进行数据共享智能合约的实现和测试。

5.1 数值仿真结果

本节首先仿真比较了本文提出的基于 Gossip 协议与传统 Gossip 协议在区块传播时延上的差异，同时分析了在验证节点比例 α 不同的情况下区块大小 s 与区块传播时延 τ 的关系；然后仿真分析了参数 ζ 和共识节点能力阈值对区块传播时延 τ 和网络去中心化程度 D 的影响；最后给出了区块传播时延 τ 和网络去中心化程度 D 的均衡化分析。本文的仿真参数设置如表 1 所示[34-35]。

表 1　　仿真参数设置

参数	值
区块大小 s /bit	2^{13}
信道容量 c /(bit · s^{-1})	100
平均往返时间 τ_{RTT} /ms	100
每个共识节点招募的验证节点所占比例 α	0.3
物联网节点计算能力最小值 σ_c /Hz	1 000
物联网节点存储能力最小值 σ_s /GB	32
共识节点的存储能力阈值 Y_c /GB	100
验证节点的存储能力阈值 Y_v /GB	60
节点总数 N /个	10 000
每比特信息需要的 CPU 周期数 U	1/64
参数 $\beta_1, \beta_2, \beta_3, \beta_4$	1, 1, 10, 10

图 4 展示了在共识节点招募的验证节点比例 α 不同的情况下区块传播时延 τ 与区块大小 s 的关系。同时将本文提出的基于 Gossip 协议的分层区块传播方案与传统 Gossip 协议的区块传播方案进行对比。其中，ζ=2，X_c =10 000 Hz，X_v = 8 000 Hz。区块大小 s 是指每个区块存储的字节数。从图 4 中可以看出，在区块大小 s 一定的情况下，相较于传统的区块传播方案，本文提出的方案区块传播时延大幅降低。在本文提出的区块传播方案中，共识节点不仅要负责共识节点层的区块传输，还要负责验证节点层的区块传输，从而使区块传播时延降低。同时，区块的传播时延 τ 随着 α 的增大而增大。因为当 α 增大时，共识节点在验证区块时招募的验证节点数增多，区块被验证的次数增多，从而导致区块的传播时延增大。在 α 一定的情况下，区块的传播时延 τ 随着区块大小 s 的增大而增大。

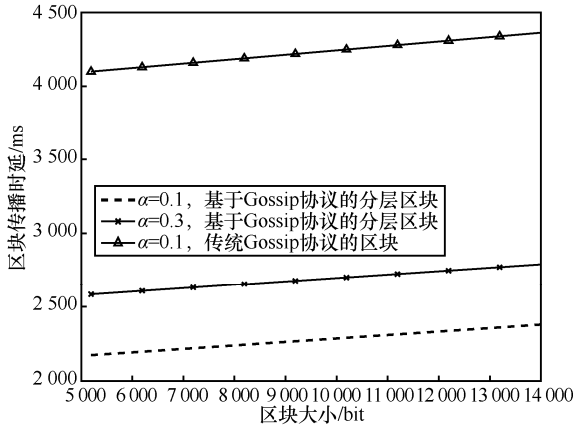


图 4 区块传播时延 τ 与区块大小 s 的关系

图 5 展示了在参数 ζ 不同的情况下区块传播时延 τ 与共识节点能力阈值关系，其中验证节点能力阈值与共识节点计算能力阈值同幅增长。从图 5(a)中可以看出，在 ζ 一定的情况下，区块的传播时延 τ 随着共识节点计算能力阈值 X_c 的增大而减小。因为当节点计算能力阈值 X_c 与 X_v 增大时，网络中共识节点以及验证节点的比例减小，区块的共识需要得到的验证次数减少，从而使区块的总传播时延减小。在 X_c 一定的情况下，区块的传播时延 τ 随着 ζ 的增大而减小。因为当 ζ 增大时，网络中共识节点和验证节点的比例减小，区块的共识需要得到的验证次数减少，从而使区块的总传播时延减小。从图 5(b)中可以看出，在 ζ 一定的情况下，区块的传播时延 τ 随着共识节点存储能力阈值 Y_c 的增大而减小。在 Y_c 一定的情况下，区块传播时延 τ 随着 ζ 的增大而减小。

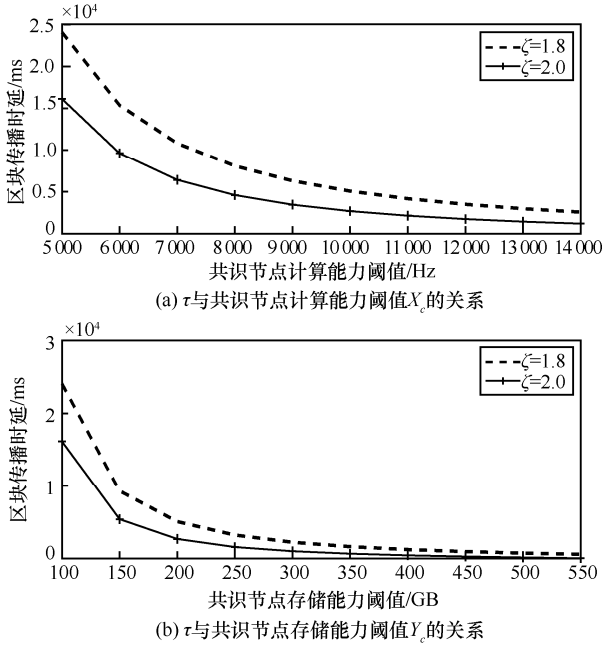
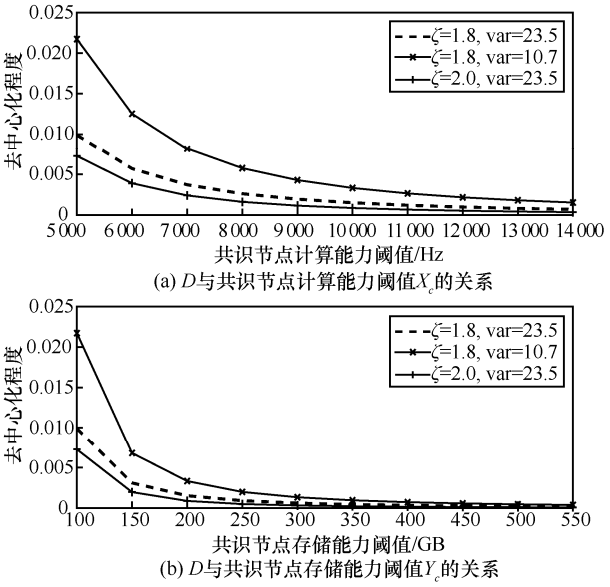


图 5　区块传播时延 τ 与共识节点能力的关系

图 6 展示了在参数 ζ 和方差 var 不同的情况下去中心化程度 D 与共识节点能力的关系，其中验证节点能力阈值与共识节点计算能力阈值同幅增长。从图 6(a)中可以看出，在 ζ 和 var 一定的情况下，区块链去中心化程度 D 随着共识节点计算能力阈值 X_c 的增大而减小。因为当节点计算能力阈值 X_c 与 X_v 增大时，区块链网络中共识节点以及验证节点的比例减小，使网络的去中心化程度 D 减小。在 X_c 和 var 一定的情况下，区块链网络去中心化程度 D 随着 ζ 的增大而减小，因为当 ζ 增大时，网络中共识

节点以及验证节点的比例减小，使网络的去中心化程度 D 减小。在 X_c 和 ζ 一定的情况下，区块链网络去中心化程度 D 随着 var 的增大而减小。从图 6(b)中可以看出，在 ζ 和 var 一定的情况下，区块链去中心化程度 D 随着共识节点存储能力阈值 Y_c 的增大而减小。



(a) D与共识节点计算能力阈值X_c的关系

(b) D与共识节点存储能力阈值Y_c的关系

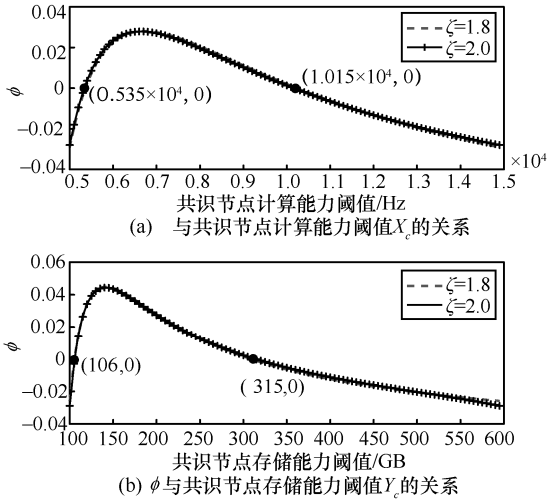
图 6 去中心化程度 D 与共识节点能力的关系

图 5 和图 6 表明，随着共识节点计算能力阈值和存储能力阈值的增加，物联网中共识节点的比例减小，区块传播时延和去中心化程度也随之减小。

图 7 展示了在参数 ζ 不同的情况下，参数 ϕ 与共识节点能力的关系。从图 7 中可以看出，ζ 对 ϕ 的影响较小。在图 7(a)中，$Y_c=100$ GB，当共识节点计算能力阈值为 0.535×10^4 Hz 或 1.015×10^4 Hz 时，$\phi\approx0$，即去中心化程度 D 与区块总的传播时延 τ 达到均衡。在图 7(b)中，$X_c=5\,000$ Hz，当共识节点存储能力阈值为 106 GB 或 315 GB 时，$\phi\approx0$，去中心化程度 D 与区块总的传播时延 τ 达到均衡。

5.2 数据共享智能合约实现

本节针对新冠肺炎疫情期间确诊患者轨迹数据共享场景，设计数据共享智能合约并进行实现。基于以太坊的 Solidity 语言以及 truffle 框架实现智能合约的编写和编译，并将智能合约部署到以太坊的 Geth 节点集群上，实现对智能合约的测试。同时，借助 IPFS 实现本文提出的共享数据地址链上存储而共享数据链下分布式存储的机制，并建立了由 3 个 Geth 节点组成的节点集群。Geth 节点硬件参数和账户信息如表 2 所示。



(a) 与共识节点计算能力阈值X_c的关系

(b) ϕ与共识节点存储能力阈值Y_c的关系

图 7 ϕ 与共识节点能力的关系

节点 1 分别与节点 2、节点 3 相互连接搭建区块链节点集群，在 geth 节点中使用 admin.peers()来获取与节点 1 相连的节点信息，包括节点 2 和节点 3 的 IP 地址、端口号和支持的协议等。当节点 1 需要发布某确诊病患的唯一标识码 patientID 和肖像数据 patientphoto 时，节点 1 通过调用智能合约中发布数据需求的函数进行信息的发布。节点 2 和节点 3 通过监听函数事件 PatientInfocreated，获得该确诊病患的信息。在监听日志中，节点 2 和节点 3 可以得到节点 1 发布的数据，其中 patientID 为 4131479，patientphoto 为确诊病患肖像的 Base64 编码在 IPFS 上的地址。

节点 2 和节点 3 收到节点 1 发布的确诊患者信息后，开始在自己的摄像头数据采集系统或者云端数据库中查找是否有该确诊病患的数据，包括确诊病患出现的时间和地点等。如果存在关于该确诊病患的数

表 2 Geth 节点硬件参数和账户信息

节点编号	节点硬件参数	节点账户地址
1	Intel i5-8520U CPU 1.8 GHz	0x9fc70af1ee9be082ed2131cfe812e41dd655c655
2	Intel i5-5200U CPU 2.2 GHz	0xc23b2a7e5a0ef861b6dc54352b0e744816e548d2
3	Intel i5-8520U CPU 1.8 GHz	0x7fe2b3ef45df5a84aec08986add17818b3471d8a

据，则将相关数据通过 IPFS 获得哈希值，然后将该数据文件的哈希值上传到区块链上。节点 2 将测试文件 test1.txt 的哈希值上传到区块链上，节点 3 将测试文件 test2.txt 的哈希值上传到区块链上，哈希值就是文件在 IPFS 网络的地址。节点 1 调用智能合约中的函数 File_Find( )，通过输入确诊病患的唯一识别码获得文件 test1.txt 和文件 test2.txt 的哈希值。节点 1 的查询结果如图 8 所示，其中，test1.txt 的存储地址为 QmYMQDn2UnoZ5vy82EzoxpkrjuBtnGXfNLRvwoYwXrbkGn，test2.txt 的存储地址为 QmcJQCNtnx1vWDvQNkvzKrDbP9Udsv5kZJXyK5jpBnYQjT。

图 8　数据需求者查询获得需要的数据文件

6　结束语

本文基于区块链技术提出一种面向 6G 的物联网数据共享和存储框架，从而实现物联网数据的去中心化共享和存储。根据物联网设备的能力，将物联网中部分节点选为共识节点和验证节点。通过 PoS 的共识机制，将区块共识和共享数据的分布式存储相结合。基于 Gossip 协议提出针对共识节点层和验证节点层的分层传播机制，同时推导了区块传播时延模型和区块链网络的去中心化评估模型。仿真分析表明，分层的传播方案相较于传统方案在区块传播时延方面有大幅降低，而且随着共识节点和验证节点的能力阈值增加，即共识节点和验证节点的比例减小，区块传播时延和区块链网络去中心化程度随之减小。当共识节点计算能力阈值为 0.535×10^4 Hz 或 1.015×10^4 Hz、共识节点存储能力阈值为106 GB 或315 GB 时，去中心化程度与区块的总传播时延达到均衡。最后，以确诊患者轨迹数据共享场景为例，借助以太坊开发平台进行数据共享智能合约的实现和测试。本文设计的数据共享和存储方案实现共享数据链下存储，即共享数据是存储在共识节点组成的分布式存储网络，而不是存储在区块链的分布式账本。数据拥有者可以根据需要控制共享数据存储的时间，从而改善了由于区块链的透明性和不可篡改性带来的数据隐私泄露问题。未来的工作将进一步探讨时变信道对区块传播时延的影响和基于区块链技术的数据共享去中心化问题。

参考文献：

[1] GIORDANI M, POLESE M, MEZZAVILLA M, et al. Toward 6G networks: use cases and technologies[J]. IEEE Communications Magazine, 2020, 58(3): 55-61.

[2] ZHOU Y Q, LIU L, WANG L, et al. Service aware 6G: an intelligent and open network based on convergence of communication, computing and caching[J]. Digital Communications and Networks, 2020, doi: 10.1016/j.dcan.2020.05.003.

[3] ZHOU Y Q, TIAN L, LIU L, et al. Fog computing enabled future mobile communication networks: a convergence of communication and computing[J]. IEEE Communications Magazine, 2019, 57(5): 20-27.

[4] YOU X H, WANG C C, HUANG J, et al. Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts[J]. SCIENCE CHINA Information Sciences, 2021, 64(1): 1-84.

[5] YANG Y, LUO X L, CHU X L, et al. Fog-enabled intelligent IoT systems[M]. Berlin: Springer, 2019.

[6] PATTAR S, BUYYA R, VENUGOPAL K R, et al. Searching for the IoT resources: fundamentals, requirements, comprehensive review and future directions[J]. IEEE Communications Surveys & Tutorials, 2018, 20(3): 2101-2132.

[7] JIANG Y N, GE X H, ZHONG Y, et al. A new small-world IoT routing mechanism based on Cayley graphs[J]. IEEE Internet of Things Journal, 2019, 6(6): 10384-10395.

[8] VERMA S, KAWAMOTO Y, FADLULLAH Z, et al. A survey on network methodologies for real-time analytics of massive IoT data and open research issues[J]. IEEE Communications Surveys & Tutorials, 2017, 19(3): 1457-1477.

[9] GE X H, ZHOU R, LI Q. 5G NFV-based tactile Internet for mission-critical IoT services[J]. IEEE Internet of Things Journal, 2019, 7(7): 6150-6163.

[10] LING L, ZHOU Y Q, YUAN J H, et al. Economically optimal MS association for multimedia content delivery in cache-enabled heterogeneous cloud radio access networks[J]. IEEE Journal on Selected Areas in Communications, 2019, 37(7): 1584-1593.

[11] YANG Y. Multi-tier computing networks for intelligent IoT[J]. Nature Electronics, 2019, 2(1): 4-5.

[12] ALI M S, VECCHIO M, PINCHEIRA M, et al. Applications of blockchains in the Internet of things: a comprehensive survey[J]. IEEE Communications Surveys & Tutorials, 2019, 21(2): 1676-1717.

[13] DANG S P, AMIN O, SHIHADA B, et al. What should 6G be?[J]. Nature Electronics, 2020, 3(1): 20-29.

[14] DAI H N, ZHENG Z B, ZHANG Y. Blockchain for Internet of things: a survey[J]. IEEE Internet of Things Journal, 2019, 6(5): 8076-8094.

[15] CHRISTIDIS K, DEVETSIKIOTIS M. Blockchains and smart contracts for the Internet of things[J]. IEEE Access, 2016, 4: 2292-2303.

[16] XIA Q, SIFAH E B, ASAMOAH K O, et al. MeDShare: trust-less medical data sharing among cloud service providers via blockchain[J]. IEEE Access, 2017, 5: 14757-14767.

[17] SHEN M, DUAN J X, ZHU L H, et al. Blockchain-based incentives for secure and collaborative data sharing in multiple clouds[J]. IEEE Journal on Selected Areas in Communications, 2020, 38(6):

1229-1241.
[18] LIU C H, LIN Q X, WEN S L. Blockchain-enabled data collection and sharing for industrial IoT with deep reinforcement learning[J]. IEEE Transactions on Industrial Informatics, 2018, 15(6): 3516-3526.
[19] SHEN M, TANG X Y, ZHU L H, et al. Privacy-preserving support vector machine training over blockchain-based encrypted IoT data in smart cities[J]. IEEE Internet of Things Journal, 2019, 6(5): 7702-7712.
[20] WANG S P, ZHANG Y L, ZHANG Y L. A blockchain-based framework for data sharing with fine-grained access control in decentralized storage systems[J]. IEEE Access, 2018, 6: 38437-38450.
[21] NGUYEN D C, PATHIRANA P N, DING M, et al. Blockchain for secure EHRS sharing of mobile cloud based e-health systems[J]. IEEE Access, 2019, 7: 66792-66806.
[22] SUN J, YAO X M, WANG S P, et al. Blockchain-based secure storage and access scheme for electronic medical records in IPFS[J]. IEEE Access, 2020, 8: 59389-59401.
[23] JIANG Y N, ZHONG Y, GE X H. Smart contract-based data commodity transactions for industrial Internet of things[J]. IEEE Access, 2019, 7: 180856-180866.
[24] YANG W L, AGHASIAN E, GARG S, et al. A survey on blockchain-based Internet service architecture: requirements, challenges, trends and future[J]. IEEE Access, 2019, 7: 75845-75872.
[25] LUU L, NARAYANAN V, ZHENG C D, et al. A secure sharding protocol for open blockchains[C]//Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM Press, 2016: 17-30.
[26] ATENIESE G, BURNS R, CURTMOLA R, et al. Provable data possession at untrusted stores[C]//Proceedings of the 14th ACM conference on Computer and Communications Security. New York: ACM Press, 2007: 598-609.
[27] DEMERS A, GREENE D, HAUSER C, et al. Epidemic algorithms for replicated database maintenance[J]. ACM SIGOPS Operating Systems Review, 1988, 22(1): 8-32.
[28] NEWMAN M E J. Power laws, pareto distributions and Zipf's law[J]. Contemporary Physics, 2005, 46(5): 323-351.
[29] SHAHSAVARI Y, ZHANG K, TALHI C. A theoretical model for block propagation analysis in bitcoin network[J]. IEEE Transactions on Engineering Management, 2020, PP(99): 1-18.
[30] AOKI Y, SHUDO K. Proximity neighbor selection in blockchain net works[C]//2019 IEEE International Conference on Blockchain. Piscataway: IEEE Press, 2019: 52-58.
[31] MISIC J, MISIC V B, CHANG X L, et al. Block delivery time in Bitcoin distribution network[C]//2019 IEEE International Conference on Communications. Piscataway: IEEE Press, 2019: 1-7.
[32] LIU X J, WANG W B, NIYATO D, et al. Evolutionary game for mining pool selection in blockchain networks[J]. IEEE Wireless Communications Letters, 2018, 7(5): 760-763.
[33] WU K, PENG B, XIE H, et al. An information entropy method to quantify the degrees of decentralization for blockchain systems[C]//2019 IEEE 9th International Conference on Electronics Information and Emergency Communication. Piscataway: IEEE Press, 2019: 1-6.
[34] KANG J W, XIONG Z H, NIYATO D, et al. Incentivizing consensus propagation in proof-of-stake based consortium blockchain networks[J]. IEEE Wireless Communications Letters, 2018, 8(1): 157-160.
[35] NIYATO D, KIM D I, KANG J W, et al. Incentivizing secure block verification by contract theory in blockchain-enabled vehicular networks[C]//2019 IEEE International Conference on Communications. Piscataway: IEEE Press, 2019: 1-7.

[作者简介]

蒋宇娜（1994− ），女，江苏徐州人，华中科技大学博士生，主要研究方向为无线通信、区块链、物联网数据共享等。

葛晓虎（1972− ），男，湖北武汉人，博士，华中科技大学教授，主要研究方向为移动通信、无线网络中的流量建模、绿色通信等。

杨旸（1974− ），男，江苏东台人，博士，上海科技大学教授，主要研究方向为 5G/6G 移动通信网络、物联网、多层次算力网络、开放无线测试验证平台等。

王承祥（1975− ），男，山东高密人，博士，东南大学移动通信国家重点实验室、紫金山实验室双聘教授，主要研究方向为无线信道测量与建模、6G 通信网络架构和关键技术、人工智能与无线通信网络的融合等。

李颉（1959− ），男，江苏苏州人，博士，上海交通大学教授，主要研究方向为大数据与人工智能、区块链、物联网、分布式系统结构与评估等。