·区块链理论与技术·　　文章编号：1000-3428(2022)06-0033-09　　文献标志码：A　　中图分类号：TP391

基于区块链的公平性联邦学习模型

陈乃月，金　一，李浥东，蔡露鑫，魏圆梦

（北京交通大学 计算机与信息技术学院，北京 100044）

摘　要：为解决典型联邦学习框架在训练样本数据分布不均衡情况下产生的聚合模型对各个客户端模型不公平的问题，结合区块链的去中心化、不可篡改性以及智能合约的特点，提出基于本地数据特征的公平性联邦学习模型，以实现数据分布差异的客户模型可信安全共享。多个客户端通过区块链上传本地参数以及信用值，利用区块链的共识机制选择信用值最高的区块进行模型聚合，在模型聚合过程中按照节点信用依次进行融合，并根据区块链记录工作节点的本地模型参数作为证据，完成整体模型参数的聚合任务，在此基础上通过广播下传当前聚合模型参数，模型利用区块链的共识机制可降低参数在传输过程中所面临的安全风险。在开源数据集上的实验结果表明，该模型相较FedAvg模型训练精度提高40%，不仅能够优化非独立同分布下的模型训练精度，同时可以防止中间参数传输信息泄露，保证了多个客户端的利益与安全隐私，从而实现具有隐私保护的公平性模型。

关键词：联邦学习；区块链；非独立同分布；公平性；隐私保护

开放科学（资源服务）标志码（OSID）：

中文引用格式：陈乃月，金一，李浥东，等.基于区块链的公平性联邦学习模型[J].计算机工程，2022，48(6)：33-41.

英文引用格式：CHEN N Y，JIN Y，LI Y D，et al.Federated learning model with fairness based on blockchain[J].Computer Engineering，2022，48(6)：33-41.

Federated Learning Model with Fairness Based on Blockchain

CHEN Naiyue，JIN Yi，LI Yidong，CAI Luxin，WEI Yuanmeng

（School of Computer and Information Technology，Beijing Jiaotong University，Beijing 100044，China）

【Abstract】This paper focuses on the unfair problem that the aggregation model generates by the non-independent and homogeneous distribution of training data in a typical Federated Learning（FL）framework.Accordingly，the characteristics of decentralization，immutability，and smart contracts of blockchain are combined to propose a blockchain-based fairness FL model to achieve trusted and secure sharing of customer models with differences in data distribution.First，each client uploads local parameters and trust values through the blockchain.Subsequently，the consensus mechanism of the blockchain was utilized to select the block with the highest trust value for model aggregation.The chain records the local model parameters of the working nodes as evidence to complete the aggregation task of the overall model parameters.Ultimately，the proposed algorithm downloads the current model parameters by broadcasting with low-security risks in the parameter transmission process.Experiments under the training of the open-source dataset reveal that the training accuracy of this model is improved by 40% compared with the FedAvg model.Accordingly，besides improving the training accuracy with the non-independent data distribution of FL，the FL algorithm also prevents information leakage in intermediate parameter transmission.Thus，it is the fairness model of privacy protection which ensures the interests，security，and privacy of each client are upheld.

【Key words】Federated Learning(FL)；blockchain；non-independent identically distribution；fairness；privacy protection

DOI：10. 19678/j. issn. 1000-3428. 0064095

0　概述

在大数据时代下，各种信息逐渐以数据的形式存在于人们的日常生活中，例如社交账号、就诊信息、浏览记录、购物记录等，很多信息会被企业收集并分析以达成预测客户行为或其他目标。但在实际中，出于保护用户数据隐私的目的，各大企业组织之间不进行数据共享，这就造成了数据通常是以孤岛的形式存在[1-3]。当前绝大多数的机器学习算法依赖大量数据，其为训练出理想的模型提供了强有力的支撑。但相关研究表明，传统机器学习模型由于存在脆弱性而容易被潜在的攻击所威胁，比如对抗

基金项目：科技创新2030—“新一代人工智能”重大项目(2021ZD0113002)。

作者简介：陈乃月(1989—)，女，讲师、博士，主研方向为联邦学习、数据挖掘；金　一、李浥东，教授、博士；蔡露鑫、魏圆梦，硕士研究生。

收稿日期：2022-03-04　　**修回日期：**2022-04-14　　**E-mail：**nychen@bjtu.edu.cn

攻击[4]、中毒攻击[5]、成员推理攻击[6]等。由于联邦学习(Federated Learning,FL)框架存在天然的数据分离以及分布式训练过程等特性,其同样容易受到不确定性攻击,因此研究联邦学习的鲁棒性问题是保证联邦学习框架能够安全稳定运行的关键。

在保证数据安全和维护数据隐私的前提下,为了解决数据孤岛的问题,研究人员提出一种新的机器学习技术,即联邦学习[7-8]。联邦学习,顾名思义就是协作学习,由多个参与方进行联合训练,其本质上是一种分布式的学习算法[9-10]。和传统的机器学习方法相比,联邦学习不需要汇集所有客户数据至中心端,在一定程度上有效防止了这些数据在上传过程中被泄露的可能性。联邦学习的主要思想是:在不共享原始数据的前提下,各个客户端利用本地数据集进行本地模型训练,完成本地训练后进行模型相关参数的交互,从而联合聚合出一个高效、共享的全局模型。但这并不意味着联邦学习方法必定是安全可信的,在聚合的过程中,外部攻击者也可以利用截获的模型相关参数重构出原始数据。并且参与客户如果出现掉线或中途退出等意外行为可能会导致整个联邦训练网络出现崩溃的情况。因此,在维护数据隐私性、解决异构性、保证公平性等方面,联邦学习仍然存在很多亟需解决的问题。

区块链作为一种去中心化的分布式存储技术具有广阔的应用前景[11-12]。不同于传统的中心化分布式存储技术,区块链中的每个节点都独立维护一份完整的数据记录,并通过共识机制实现数据存储的一致性,当单个节点发生故障或者被恶意攻击时,不会影响其他节点的正常运行,从而保证了数据的安全性,提高了整个系统的鲁棒性[13]。区块链中每个节点按照时间顺序相连的区块存储数据,每个区块的区块头中记录着上一个区块的哈希值,某个区块的改变将会影响后面一系列区块的计算,所以区块链上伪造或者更改数据的代价很高,遭受恶意节点篡改全部数据的可能性极小。此外,区块链还通过非对称加密算法、数字签名等密码学技术进一步保证了数据的安全。因此,区块链技术可以满足保护数据安全和用户隐私的要求,实现多个参与方之间的可信数据交换。结合联邦学习的架构,区块链技术可以激励对聚合全局模型贡献度高的用户更加积极地参与训练,形成整体可持续的学习模式。同时,利用共识机制和激励机制可以更加公平地计算并分配参与方的收益,实现联邦学习各个节点间贡献度的公平性。因此,利用区块链技术可以识别出联邦训练过程中可能存在的恶意攻击行为或潜在危险,追溯恶意用户并进行惩罚。

目前,结合联邦学习和区块链的研究工作已经取得了较好的进展。文献[14]提出一种结合区块链和联邦学习实现安全数据分享的框架,通过许可链建立多个参与方之间的安全连接,可以进一步控制节点对共享数据的访问,降低了联邦学习协作过程中数据泄露的风险。文献[15]使用区块链取代联邦学习中的中央服务器,使得参与节点在无可信第三方的情况下依然可以建立可信通信,并对该框架在效率、隐私保护、抗中毒攻击等方面的性能进行了研究。DeepChain[16]对不信任的参与节点提供基于区块链价值驱动的激励机制,并且通过区块链实现联邦学习协作过程的可审计性,对不诚实的节点进行惩罚,对诚实的节点进行奖励。以上研究在区块链的基础上,对联邦学习协作过程中的数据安全、隐私保护和公平性等方面的问题进行了改善,但是没有考虑到不同参与节点数据分布不均衡问题对联邦学习协作过程的影响。文献[17]则提出一种多层分布式防御计算框架,有效改善了本地节点训练数据有限的问题,但是模型结构较复杂,在应用上存在较大的困难。

针对客户端训练样本分布不均衡现象导致的训练精度问题,本文结合区块链的共识技术研究高效的公平性联邦学习机制。首先根据客户端本地训练模型参数计算节点可信度,构建基于信任度的联邦学习聚合模型,各客户端结合本地模型的历史性能更新本地模型参数,以保证本地数据特征不被全局模型遗忘。然后利用区块链的共识机制技术保障联邦学习中各参与方训练数据传输的可信性,确保节点间数据传输的隐私性及不可篡改性。最后通过在各客户端利用不同数据分布的训练集进行实验仿真,以保证各客户端的隐私性。

1 基础内容和相关研究

1.1 基础内容

联邦学习是一种由多方共同参与进行联合学习的新型技术。联邦学习结构模型如图1所示,主要包括拥有数据的客户端和包含聚合模型的中央服务器两个主体。相比一般的分布式机器学习,存在的明显不同是联邦学习不需要将数据集进行聚合操作,大幅降低了在集合多方数据的阶段造成隐私泄露问题的可能性。在每轮迭代中,各个参与方持有本地数据进行本地模型训练,并将更新后的模型相关参数值上传至中央服务器,由服务器聚合参数,从而更新全局模型。

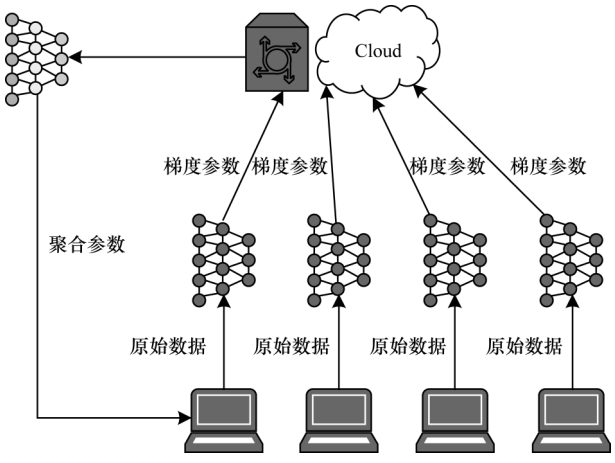

图1 联邦学习结构模型

Fig.1 Federated learning structure model

作为一种分布式数据库技术，区块链可以在多个互不了解的参与方之间实现可信的数据共享[18]。区块链是去中心化的，无需借助可信的第三方，由多个参与方按照严格的规则共同维护，通过密码学、共识机制、智能合约等多种技术实现交易过程的高度可信性，可以有效减少单个故障节点或恶意节点带来的影响。

区块链通过以区块为单位的数据结构保障了数据的不可篡改性，每个区块包括区块头和区块体两部分，由区块头连接上一个区块，由区块体记录交易数据。在区块链系统中，每个参与节点都会记录所有的交易信息，并对其他节点记录的信息进行正确性验证。存储在区块链中的交易信息是公开透明的，同时交易账户的个人信息是高度加密的，这样既提高了数据库的安全性，又保证了参与方的隐私性，还可以有效追溯非法篡改的恶意节点，对其进行惩罚。

1.2　国内外研究现状

谷歌提出关于联邦学习的算法，使用该算法来处理安卓手机终端的本地模型更新。近年来由联邦学习结合其他算法有效地处理了一系列实际问题，而其中如何提高通信开销、解决异构性、保证公平性和隐私性等问题都是当前联邦学习算法主要的研究方向。联邦平均(FedAvg)[10]算法是由MCMAHAN等提出的，和传统FedSGD算法相比，在一定程度上降低了通信成本，但是其存在所有客户共用相同模型的问题，无法满足持有非独立同分布数据客户端的需求。为了解决不同客户端设备状况存在差异、所处网络环境存在不同以及非独立同分布数据等问题，文献[19]在联邦平均的基础上提出了一种异步联邦优化算法处理设备异构性，文献[20]则针对模型异构性将联邦学习与元学习相融合。在FedAvg算法的基础上添加L2正则化项而构造的FedProx模型[21]，是通过添加约束项使得客户端局部更新过程中的参数尽可能与服务端模型的参数相近。通过此简单且有效的操作，使得模型可以有效缓解数据异质性的问题，同时能够安全地聚合来自不同客户端具有模型差异性的梯度信息进而提高模型的预测性能。文献[22]通过利用一阶模型优化来计算模型聚合的权重，客户端可以根据其与其他客户端的权重系数进行加权聚合，因而不同的客户端可以拥有不同的全局个性化模型，进而利用合理的权重信息来得到更加精确的全局模型。然而，这些联邦学习优化方法只考虑了客户端在本轮模型训练的性能，在节点选择过程中容易造成历史模型的遗忘，本文方法将根据历史模型个性化更新本地模型，以更实际地体现整体模型与本地模型的关系。

区块链技术最早由中本聪于2008年提出，并在2009年得到实践。区块链的发展经历了以数字货币为特点的1.0时代、以智能合约为特点的2.0时代与正在迈向更加安全和完备的3.0时代。区块链通过块链式数据结构存储数据，根据分布式共识算法更新数据，通过智能合约操作数据。目前，区块链的共识算法主要包括工作量证明、权益证明、委托权益证明和实用拜占庭容错等。文献[23]提出优化算法T-PBFT，通过团队签名和相互监督选择信用较好的节点参与到共识过程中，提高系统的容错性。文献[24]提出一种改进的委托权益证明共识机制，引入信用机制，通过监测节点行为及时剔除不积极参与协作的恶意节点。

目前，很多联邦学习算法在处理一些实际应用时得到的模型效果存在一定局限，只有当问题的数据类型为独立同分布或数据规模较小时，这些算法能够实现较为理想的结果。

然而，在实际问题中面对的情况往往更加复杂，一般的联邦学习技术缺乏一定的可行性，不能有效地平衡例如隐私保护、公平性等一系列重要因素。因此，很多研究的方向是将联邦学习与其他技术相结合来优化训练模型的性能，其中，将其与区块链技术进行融合是一个热门方向。现有的一些区块链结合联邦学习的算法主要动机是通过激励更多高质量的数据拥有者参与到协作学习过程中以提升模型的训练效果。区块链中有很多设计成熟的激励机制，包括增加收益型激励机制、赋予权利型激励机制和提高声誉型激励机制等类型。文献[25]提出一种基于样本数目大小的激励机制，给予区块链中矿工与其关联设备训练中使用数据量成正比的奖励，但在该机制中，恶意节点可能通过造假骗取奖励。文献[26]则通过引入与数据质量相关的参数，不仅考虑到样本数据量的大小，还考虑到样本数据对联邦学习训练过程准确率的影响，激励更多拥有高质量数据的用户参与协作学习。文献[27]则使用声誉作为矿工的选择标准，通过主观逻辑对矿工进行声誉评价，并将声誉评价公开透明地记录在区块链中，用于可信的声誉计算，实现联邦学习过程中可靠参与者的选择。

本文与现有的基于区块链的联邦学习工作相比，主要贡献为采用基于损失函数的梯度投影算法保证数据分布不均衡情况下各个用户的数据特征，同时，采用结合本地模型的历史性能更新本地模型参数，以保证本地数据特征的不被全局模型遗忘。并且，通过区块链技术来保证节点间数据传输的隐私性以及不可篡改性。同时，设计了PoT共识机制作为联邦学习聚合的引擎，适用于轻量级智能边缘计算中的联邦学习。该方法能够降低模型中心化参数泄露的风险，降低区块链节点的通信开销。

2　系统模型分析

2.1　本地模型训练

相较于传统的分布式机器学习，联邦学习模型可以保证网络数据高效处理，中心节点不需要训练网络中所有的数据，而是将各个参与节点上传的参数进行综合处理，进而返回给各个节点相应的学习参数。

分布式机器学习利用多个计算节点进行机器学习或深度学习，提供扩展大规模训练数据的计算和存储能力。其中，训练数据被分为不相交的数据分片并被发送给各个工作节点，工作节点在本地执行随机梯度下降后，将模型参数返回给中心服务器。由于在该过程中，中心节点不会接触到各个节点的

原始数据,因此能够保证各个节点本地数据的隐私性。同时,由于中心节点只需处理边缘节点上传的相关参数,因此降低了中心节点的计算量,提高了机器学习的工作效率。

本文将结合联邦学习框架和区块链的思想设计公平性的联邦学习模型,采用信任机制量化工作节点的本地模型性能,通过区块链传递各个节点模型参数,其框架如图2所示。

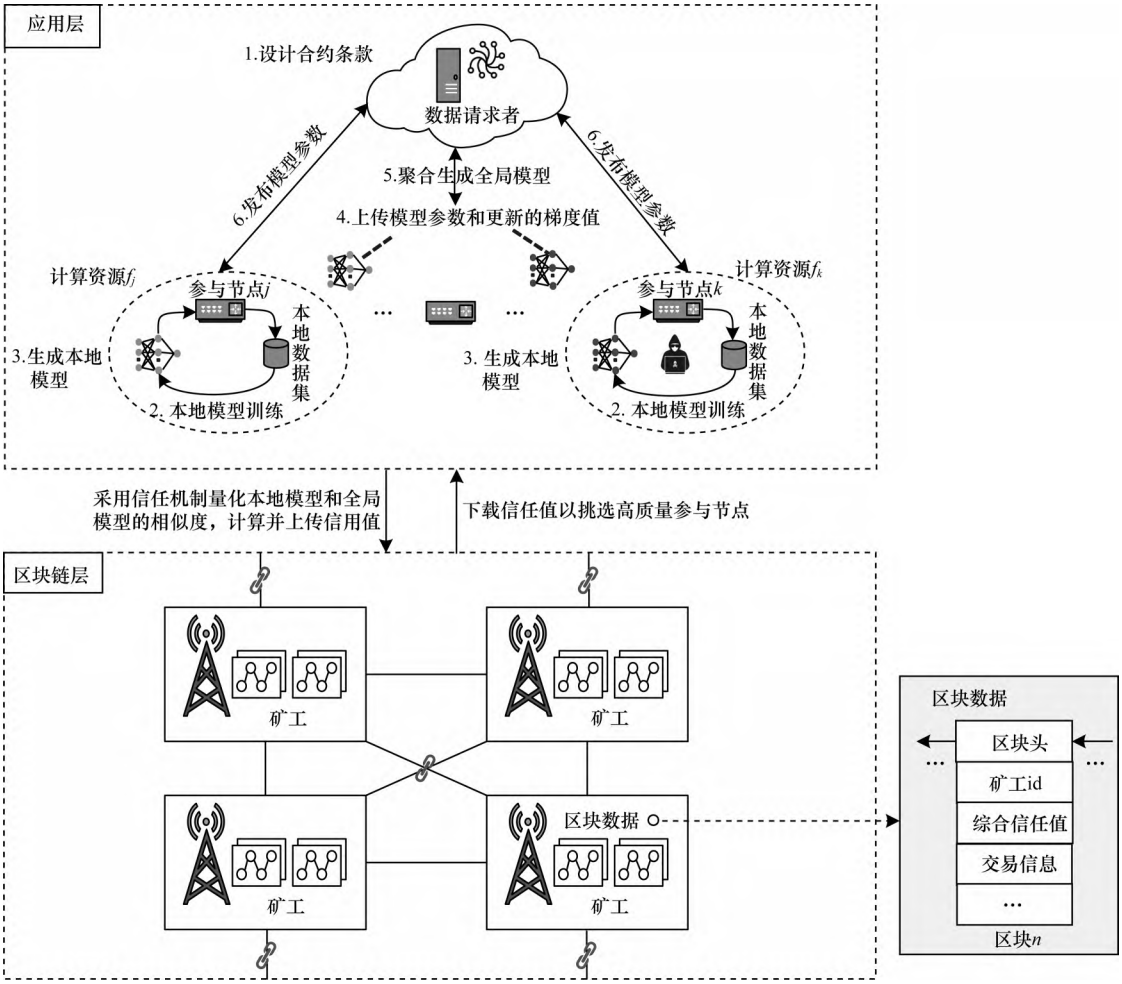

图2 系统模型框架

Fig.2 System model framework

在联邦学习框架中,参与工作的节点定义为$M=\{m_1,m_2,\cdots,m_k\}$,其中每个工作节点训练本地机器学习模型,其模型参数为$W=\{w_1,w_2,\cdots,w_k\}$。在本地模型中采用成熟的神经网络进行数据特征的分类,本文采用卷积神经网络方法构建神经网络模型,提取原始数据中最有效的特征表示。该网络的构建包括卷积层、池化层和全连接层。如图3所示,该模型采用两个卷积层和池化层交叠,然后连接全连接层。

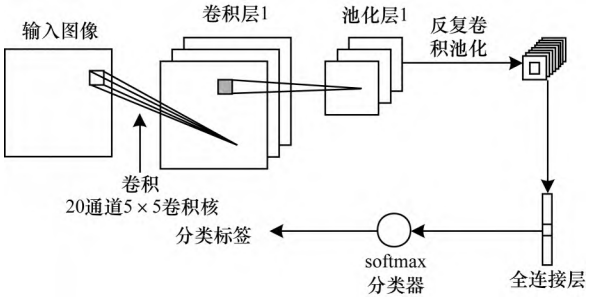

图3 本地CNN模型

Fig.3 Local CNN model

卷积层主要将网络连接和系统审计数据转化为特征图,有效提取数据特征。在卷积层中,随机初始化其权重w和偏置值b。卷积计算以滑动窗口的形式进行,计算公式如下:

$$c_i=H(wx_i+b) \tag{1}$$

其中:x_i表示原始数据;c_i表示计算后的数据特征。

池化层阶段主要进行降采样操作,降低特征图的特征空间,否则过多的特征图参数不利于高层特征的抽取。特征的降维操作一般采用最大池和平均池的两种计算方式。

$$c_{\max}=\max(c_1,c_2,\cdots,c_i) \tag{2}$$

$$c_{\text{ave}}=\frac{\sum c_i}{i} \tag{3}$$

其中:最大池化是保留每个小区域中的最大值,即重点在于该区域是否匹配上,而非具体某一处的高度匹配。

在全连接层,将根据前两层进行的特征提取进行数据特征的分类操作。该层类似于前馈神经网络

中的隐含层，不再以空间拓扑结构进行计算，而是将其通过激励函数展开为向量的形式。

2.2　PoT共识算法

由于区块链具有分布式记账的特点，可以保证区块链网络中联邦学习的参与方之间交互是安全可靠的，且共享的数据具备透明、一致、防篡改等特性。以区块链为基础的联邦学习框架需要通过共识机制来保证参与联邦学习的各个节点进行可信的协作学习。在该框架中利用区块链来接收和保存参与节点相关的认证信息和模型参数，再经由共识协议对其进行认证处理，有效保证了协作训练的公平性。

目前，一般的区块链共识机制并不能够满足联邦学习在通信开销和公平性方面的需求。区块链中最基本的共识机制包括工作量证明（Proof of Work，PoW）[26-27]共识机制和权益证明（Proof of Stake，PoS）[28-29]共识机制。虽然PoW的原理更简单，由于该机制选择的依据是节点的算力竞争，需要被证明人执行大量重复的计算工作，容易导致资源严重损耗，达成共识的周期较长。PoS能够有效缓解资源浪费的问题，达成共识的时间较短，但是需要被证明人拥有一定数量的加密货币所有权，安全性能低，拥有加密货币多的用户容易被攻击。因此，本文对共识机制进行适当改进，提出一种更适合联邦学习应用场景的基于信任值证明（Proof of Trust，PoT）的共识机制，引入了信任值来衡量参与节点在联邦学习协作过程中的表现情况。

PoT通过利用每个节点在联邦学习协作过程中的可信度来达成节点间的共识，且实现的去中心化程度和安全性能较高，既可以保证较低的通信开销，也能保证每个参与节点在协作过程中公平地获益，从而鼓励更多高质量、可信任的节点积极参与到协作过程中。在PoT中，每个节点都按照共识规则进行数据处理，使得整个分布式系统中所有节点保持一致性。节点通过选择性地参与到联邦学习协作过程中来更新信任值，在一段时间内，具有最高信任值的节点被选取作为主节点，获取唯一记账权，生成新区块时需要节点提供一个有效的信任值证明，并且这个证明应可以被其他协作节点验证。通过对节点在联邦学习协作过程中的表现进行认证来证明该节点拥有一定的可信度，增大了节点作假的成本，在一定程度上保证了整个网络的安全性。因此，PoT解决了PoW中资源损耗严重和PoS中安全性能低的问题，可以更好地运用于联邦学习应用场景。

在联邦学习协作过程中，如果恶意节点上传虚假的本地参数，会导致最终聚合的全局模型质量受到影响。为了防止恶意节点破坏协作过程，本文通过信任值来衡量每个节点上传的本地模型参数的可信度。本文中计算信任值的主要依据是本地模型和全局模型的相似程度，定义如下：

$$s_i^{(t)}=\alpha s_i^{(t-1)}+(1-\alpha)\tilde{s}_i^{(t)} \tag{4}$$

其中：$\tilde{s}_i^{(t)}=\cos(\Delta w_g^{t-1},\Delta w_i^t)$，$\Delta w_i$是工作节点$i$本地模型更新的参数，$\Delta w_g$为当前轮次的全局模型参数；$\alpha$为历史系数协调当前轮次与历史模型的比例。在历史信任值的基础上，节点通过每轮次训练得到的本地模型参数与上轮次中聚合得到的全局模型参数之间的相似程度更新信任值，基于历史的信任值可以更好地衡量节点在联邦学习协作过程中的整体表现情况。

在每一轮训练中，当随机选择的所有参与节点完成本地训练后，各个节点将在区块链网络中共享透明的、不能随意篡改的全局模型参数，随后进行共识验证。在共识过程中，首先根据参与本轮训练的节点组在前几轮的整体训练效果选出信任值最高的节点P_{leader}，即主节点，并由主节点处理交易请求。当接收到一条交易请求时，该节点根据拥有的本地数据和模型参数进行训练，将训练得到的损失$\text{loss}(f(x_i),y_i)$通过广播的形式发送给区块链网络上的其他参与节点p_j，同时收集相关的交易信息并包装成一个区块Block_t进行广播。由主节点主持的新区块主要包含两部分内容，即$\text{Block}_t=\{<H_{t-1},H_m>,T[w_i,\text{loss}(f(x_i),y_i)],s_i^{(t)}\}$。其中：区块头中存储前一区块的哈希值$H_{t-1}$和本区块相关哈希值$H_m$；主节点的信任值证明$s_i^{(t)}$以及主要的交易信息$T$则保存在区块体中。

除主节点外的其他参与节点p_i对该区块Block_t进行共识验证，只有当广播的区块通过这些节点的验证，才能证实主节点发布的内容是可信的。在共识节点p_j验证区块的过程中，首先需要验证存储在区块头中的哈希值确保区块发布者不可随意篡改区块内容，其次验证主节点的数字签名和其他信息来认证节点身份。此外，这些共识节点利用本地数据对接收到的模型参数w_i以及$\text{loss}(f(x_i),y_i)$验证主节点的信任值$s_i^{(t)}$的可信程度，并将最终的验证结果进行广播。综合其他节点对该区块的验证结果，最终认证为合法的区块Block_t则由主节点添加到对应的区块链中。

2.3　联邦模型成员审计

在区块链网络中，可能存在某些参与节点通过广播造假其模型参数以恶意影响全局模型更新的效果。为了避免这种情况的发生，本文提出一种由参与节点共同制约不同节点的本地模型质量评估的联邦模型审计算法。在该算法中，当节点完成一次迭代训练就会将其训练更新的本地模型参数广播给其他参与节点，也就是协作节点，由协作节点对该模型参数进行质量评估。其中，评估的方法是协作方基于本地数据集对模型参数进行测试，最终以计算得到的模型误差作为评价结果并广播至链上。不同于简单地以节点自身的模型训练损失作为衡量本地模型质量的唯一评价指标的方法，联邦模型成员审计通过综合考虑节点的直接评价和其他协作节点的间

接评价，可以得到一个更加公正客观的模型质量评价结果。

节点本地模型质量评估的联邦模型审计算法通过区块链上的交易完成。参与联邦学习训练过程的节点在完成本地训练之后，得到本地模型参数w_i和训练损失$\mathrm{loss}(w_i)$，然后发送交易请求，将w_i、$\mathrm{loss}(w_i)$等交易数据发送给其他参与协作的节点。为了保证交易过程的安全性，本文通过使用数字签名算法确保交易信息的来源安全可靠，并通过使用非对称加密算法确保交易数据的安全。发送节点使用接收节点的公钥对交易数据进行加密，同时使用自己的私钥进行数字签名，最后将加密的交易数据、数字签名和自己的公钥一起发送出去。接收节点在接收到交易信息后，通过发送节点的公钥对数字签名进行验证，在确认数据未被篡改且交易有效之后，通过自己的私钥对交易数据进行解密，最后根据所拥有的本地数据在收到的模型上进行训练，得到训练误差$\mathrm{loss}(w_i)$，同样通过交易的方式将$\mathrm{loss}(w_i)$广播给参与到联邦学习训练过程的其他节点。被验证模型的质量评价结果综合了所有节点测试该模型得到的训练误差，质量评价结果的计算公式如下：

$$\mathrm{loss}^t(w_i)=\mathrm{loss}(w_i)+\frac{1}{n-1}\sum_j \mathrm{loss}_j(w_i) \tag{5}$$

在执行PoT共识的阶段，为了实现更加公平可信的联邦协作学习，当主节点$P_{\text{;eader}}$将区块Block_t广播给其他参与节点p_j后，这些训练协作节点测试接收到的模型参数w_i，计算出对应的模型损失作为审计结果。这种方法大幅降低了节点上传造假的模型损失至区块链且不被发现的可能性。并且，由于具有独特的不可篡改的分布式账本特性，区块链可以实现联邦学习协作过程的全程可追溯，参与节点在训练过程中产生的数据记录被永久储存且不可伪造，从而有效预防不诚实的节点上传低质量的模型参数而危害整体模型聚合。

2.4 本地更新算法

通过PoT共识机制选出主节点，区块链上其他节点收集到参与训练节点的本地模型后进行模型聚合，现有的模型聚合方法主要是联邦平均算法，但是联邦平均算法没有考虑各个节点间数据的异质性以及样本标签的不均衡问题，这会导致标签少的样本的特征提取不公平。

在区块链选择工作节点的过程中，不能确保每一轮的工作节点都具有相同的数据分布、数据规模等特征。每个本地区块都有本地的数据特征，特别是当数据是非独立同分布时，不同本地模型之间参数模型的差异就会增大。这种现象会造成在联邦学习框架中产生梯度的冲突，仅通过对各个区块参数的加权平均可能会消除梯度方向间的差异，造成梯度方向差别信息的遗漏。为了缓解节点间样本标签不均衡问题，本文在聚合本地模型参数前，计算梯度之间是否存在很明显的方向冲突，进而采用梯度投影的思想聚合全局模型。其中的模型梯度冲突一方面来源于拥有相同特征的数据客户在某轮通信中占据多数，导致与其他特征用户模型梯度产生明显差异，另一种可能性来源于训练的数据量不均衡所导致的模型梯度之间的差异。因此，本文在本地模型参数出现过大差异时，选择采用按序投影的方式以保留各个用户的数据特征。由于在梯度投影过程中最后被投影的参数所保留的方向最全面，因此本文采用按照本地训练损失函数的逆序进行全局冲突的梯度投影。

$$w_{\mathrm{g}}^t=\frac{1}{n}\left[\sum_i s_i^{t-1}w_i^{t-1}-\sum_i^m\left\|w_i^{t-1}\right\|\cos\left\langle w_i^{t-1},w_j^{t-1}\right\rangle\frac{w_j^{t-1}}{\left\|w_j^{t-1}\right\|}\right] \tag{6}$$

同时，为了使得本地更新模型可以更好地适合其自身数据集的特征，在模型每个区块中节点更新本地模型参数时，将对比上一轮全局模型的性能和本次本地模型的性能，以自适应地调整更适配本地数据集的模型权重。本文将结合本地模型的历史性能优化全局模型在区块链中各个节点的模型训练参数。

$$w_i^{t+1}=(1-\beta)w_{\mathrm{g}}^t+\beta w_i^t \tag{7}$$

其中：w_{g}^t表示在t时刻的全局模型参数；w_i^t表示在t时刻的本地节点i的模型参数；β表示在本地模型更新中本地特征历史参数所占比例，该比例通过两次模型训练的表现得到本轮所占的比例。

$$\beta=\frac{\left|\Delta\mathrm{loss}_i^t\right|}{\mathrm{loss}_{\mathrm{g}}^{t-1}} \tag{8}$$

当上一轮的全局模型在本地的表现相较于本地模型较差时，可能是由于本地数据分布的特征与全局模型聚合中大部分的分布特征不相同，此时需要考虑本地的数据特征，因此，在本地模型更新过程中，将会调整全局模型所占比例，强调本地模型所表现的数据特征，以更好地优化本地模型更新，其模型聚合更新算法描述如算法1所示。

算法1 模型聚合更新算法

输入 $\mathrm{loss}(w_i)$，w_i，s_i

输出 w_{g}

1. 按照$\mathrm{loss}(\mathrm{w_i})$升序进行$\mathrm{w_i}$排列，形成$\mathrm{Rank}\{\mathrm{w_i}\}$

2. 对于每一个节点根据其在$\mathrm{Rank}\{\mathrm{w_i}\}$中的顺序依次进行梯度投影

3. 主节点计算

$$\mathrm{w_g^t}=\frac{1}{\mathrm{n}}\left[\sum_{\mathrm{i}}\mathrm{s_i^{t-1}}\,\mathrm{w_i^{t-1}}-\sum_{\mathrm{i}}^{\mathrm{m}}\left\|\mathrm{w_i^{t-1}}\right\|\cos\left\langle\mathrm{w_i^{t-1}},\mathrm{w_j^{t-1}}\right\rangle\frac{\mathrm{w_j^{t-1}}}{\left\|\mathrm{w_j^{t-1}}\right\|}\right]$$

4. 广播给其他区块链上节点

5. 各个区块计算该轮$\mathrm{w_g}$的性能表现，并更新本地模型参数$\mathrm{w_i^{t+1}}=(1-\beta)\mathrm{w_g^t}+\beta\mathrm{w_i^t}$

6. 返回$\mathrm{w_g^t}$

3　实验与结果分析

为了验证本文算法的有效性，采用机器学习中图像识别任务中被广泛使用的MNIST数据集和CIFAR-10数据集进行实验，与联邦平均算法进行异常检测精度和数据处理效率的对比。MNIST数据集由0到9手写数字图片和数字标签所组成，每个样本都是28×28像素的灰度手写图片，共计60 000个训练样本和10 000个测试样本。相较于MNIST数据集，CIFAR-10数据集是3通道的彩色RGB图像，且图片尺寸为32×32像素，比MNIST稍大。相比于手写字符，CIFAR-10含有的是现实世界中真实的物体，不仅噪声很大，而且物体的比例、特征都不尽相同，这为识别带来很大困难。本文实验代码基于Python、PyTorch实现联邦学习协作式模型训练，通过线程池模拟多节点分布式并行训练。

3.1　实验环境及测试标准

本文实验计算机配置为16 GB内存，i7-9700 3.2 GHz处理器，操作系统为Centos，实验语言为Python，框架搭建使用Pytorch平台完成。本文使用查全率（AC）对模型进行评估。计算公式如式（9）所示：

$$A_{AC}=\frac{T_{TP}+T_{TN}}{T_{TP}+F_{FN}+F_{FP}+T_{TN}} \tag{9}$$

其中：T_{TP}表示正确分类的正样本数，即预测为正样本，实际也是正样本；F_{FP}表示被错误地标记为正样本的负样本数，即实际为负样本而被预测为正样本，所以是False；T_{TN}指正确分类的负样本数，即预测为负样本，实际也是负样本；F_{FN}指被错误地标记为负样本的正样本数，即实际为正样本而被预测为负样本，所以是False；$T_{TP}+F_{FN}+F_{FP}+T_{TN}$表示样本总数；$T_{TP}+F_{FN}$表示实际正样本数；$T_{TP}+F_{FP}$表示预测结果为正样本的总数，包括预测是正确的和错误的；$F_{FP}+T_{TN}$表示实际负样本数；$T_{TN}+F_{FN}$表示预测结果为负样本的总数，包括预测是正确的和错误的。

3.2　结果分析

本文实验主要有3个方面：首先分析在独立同分布情况下该算法与对比算法的性能差别；其次面向非独立同分布数据特征进行模型训练，更贴近真实场景下联邦学习的各个节点数据分布特征，从而验证本文算法的公平性，对于不同的节点可以公平地满足各自的特征需求；最后验证区块链模型聚合的服务性能，保证该算法能够在区块链框架下高效运行。

为了证明提出算法的有效性，该实验将MNIST数据集和CIFAR-10数据集按照数据集规格分割为同等比例的10份数据集分给各个工作节点，其中每个数据集中包含各类照片数据。通过这种数据分割方式可以使得各个节点之间的数据达到独立同分布的状态。将本文算法与原始联邦学习算法作为比较基准，结果如图4、图5所示。

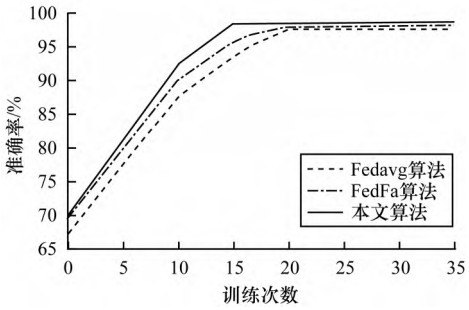

图4　MNIST数据集中不同算法性能对比

Fig.4　Performance comparison of different algorithms in MNIST dataset

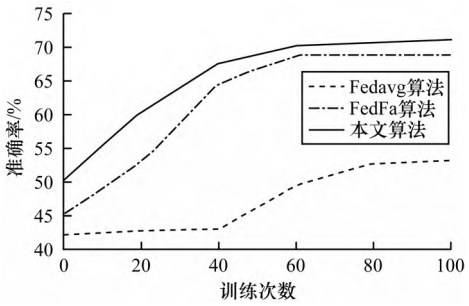

图5　CIFAR-10数据集中不同算法性能对比

Fig.5　Performance comparison of different algorithms in CIFAR-10 dataset

如图4所示，在MNIST数据集中各算法的表现均较好，但是本文算法无论在初始状态还是随轮次的迭代，均高于Fedavg算法。并且该算法在迭代15次达到了收敛状态，相较于Fedavg算法和FedFa算法都是最先达到收敛状态的。如图5所示，本文算法在训练CIFAR-10数据集时，迭代60轮次达到收敛，且准确率达到70.2%，FedFa算法训练65轮次直至收敛，且准确率为68.7%。因此，通过与baseline和基于公平性的联邦学习方法对比，本文算法可以在较少的迭代次序中得到更高的准确率，减少了通信次数。

为了分析该算的公平性，本次实验将MNIST数据集按照标签进行预处理，使得每个工作节点仅仅存在部分标签数据，而不包含全部的10类标签以符合实际工作中数据分布不均衡的现象。

通过图6可以发现，随着通信轮次的增加，模型的准确率也在提高，本文算法在开始几轮通信中并没有显著的优势，这是由于开始时模型参数的随机性较大，但是很快就实现了准确率的提升，并且提升的速度较于Fedavg算法有很大的进步。Fedfv算法采用梯度投影方法缓和梯度冲突造成的影响，该算法通过设计不同客户端的投影顺序以优化整体模型准确率，其准确率随着迭代次数的增加而提高。但是该算法没有考虑各个节点的历史表现性能，因此，本文提出的算法随着模型训练的迭代，其性能均优于其他3种算法，且较快地达到模型的收敛状态。因此，可以证明在节省通信成本和完成非独立同分布情况下，本文算法在分类任务方面具有一定的优势。

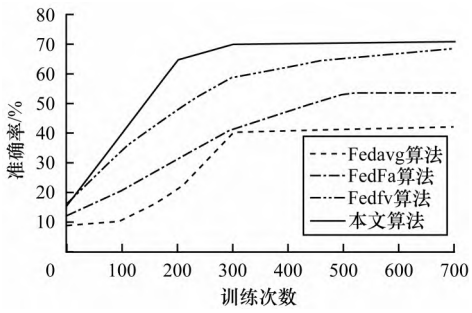

图6 non-IID数据集中不同算法性能对比

Fig.6 Performance comparison of different algorithms in non-IID dataset

本文实验采用Python Flask框架实现区块链服务功能，能够提供创建交易、链状态查询、节点注册等服务。通过区块链网络构建搭载联邦学习中的工作节点，假设每个节点存储10次联邦学习模型聚合和更新的交易，本次实验在MNIST数据集中分析基于区块链公平性联邦学习模型在区块链中的吞吐量和时延性能。

如图7所示，本实验分析在不同交易数量情况下的吞吐量与时延的比例，可以发现随着发送交易数量的增加，交易吞吐量的也随之增加，在2 000以内的交易数量中吞吐量稳步线性增长。但是随着交易数量的再度增大，其吞吐量的增长速度变得缓慢，这是由于在预备块的生成和区块达成共识时需要保持一定数量节点的一致性，从而产生较高时延以产生吞吐量增速下降的现象。本文进行30次实验取得均值以展示交易吞吐量相对稳定的状态。因此，通过区块链结构同时打包区块信任度以及模型参数进行区块交易是可行的，并且能够满足一定的性能需求。

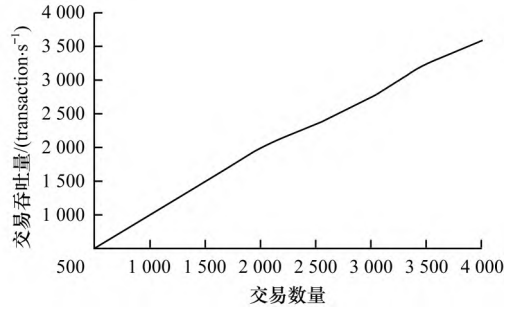

图7 区块链的交易吞吐量

Fig.7 Transaction throughput of blockchain

4 结束语

本文主要研究在数据分布不均衡情况下数据共享模型公平性问题，通过结合联邦学习和区块链技术，提出基于本地数据特征的公平性联邦学习模型。设计基于信任度的共识机制，利用区块链记录工作节点的本地模型参数作为证据实现模型聚合功能。实验结果表明，本文算法可以优化非独立同分布下的模型训练精度，防止中间参数传输信息泄露，实现区块链在公平性的联邦学习框架中的应用。下一步将扩展本文节点信任机制，研究更公平的整体模型聚合方法及加入隐私保护的梯度更新方法，以降低隐私泄露的风险。

参考文献

[1] LIANG Y, GUO Y G, GONG Y X, et al. An isolated data island benchmark suite for federated learning[EB/OL]. [2022-02-01]. https://arxiv.org/abs/2008.07257.

[2] YANG Q, LIU Y, CHEN T J, et al. Federated machine learning: concept and applications[J]. ACM Transactions on Intelligent Systems and Technology, 2019, 10(2): 12.

[3] LIM W Y B, LUONG N C, HOANG D T, et al. Federated learning in mobile edge networks: a comprehensive survey[J]. IEEE Communications Surveys & Tutorials, 2020, 22(3): 2031-2063.

[4] GOODFELLOW I J, SHLENS J, SZEGEDY C. Explaining and harnessing adversarial examples[EB/OL]. [2022-02-01]. https://arxiv preprint arxiv: 1412.6572.

[5] XIAO H, BIGGIO B, BROWN G, et al. Is feature selection secure against training data poisoning?[C]//Proceedings of IEEE International Conference on Machine Learning. Washington D. C., USA: IEEE Press, 2015: 1689-1698.

[6] SHOKRI R, STRONATI M, SONG C Z, et al. Membership inference attacks against machine learning models[C]// Proceedings of IEEE Symposium on Security and Privacy. Washington D. C., USA: IEEE Press, 2017: 3-18.

[7] POKHREL S R, CHOI J. Federated learning with blockchain for autonomous vehicles: analysis and design challenges[J]. IEEE Transactions on Communications, 2020, 68(8): 4734-4746.

[8] ZHANG C, XIE Y, BAI H, et al. A survey on federated learning[J]. Knowledge-Based Systems, 2021, 216: 106775.

[9] MOTHUKURI V, PARIZI R M, POURIYEH S, et al. A survey on security and privacy of federated learning[J]. Future Generation Computer Systems, 2021, 115: 619-640.

[10] MCMAHAN B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data [C]//Proceedings of PMLR' 17. Washington D. C., USA: IEEE Press, 2017: 1273-1282.

[11] 张亮，刘百祥，张如意，等. 区块链技术综述[J]. 计算机工程，2019，45(5)：1-12.
ZHANG L, LIU B X, ZHANG R Y, et al. Overview of blockchain technology[J]. Computer Engineering, 2019, 45(5): 1-12. (in Chinese)

[12] HAKAK S, KHAN W Z, GILKAR G A, et al. Securing smart cities through blockchain technology: architecture, requirements, and challenges[J]. IEEE Network, 2020, 34(1): 8-14.

[13] FAROUK A, ALAHMADI A, GHOSE S, et al. Blockchain platform for industrial healthcare: vision and future opportunities[J]. Computer Communications, 2020, 154: 223-235.

[14] LU Y L, HUANG X H, DAI Y Y, et al. Blockchain and federated learning for privacy-preserved data sharing in industrial IoT [J]. IEEE Transactions on Industrial Informatics, 2020, 16(6): 4177-4186.
[15] QU Y Y, GAO L X, LUAN T H, et al. Decentralized privacy using blockchain-enabled federated learning in fog computing[J]. IEEE Internet of Things Journal, 2020, 7(6): 5171-5183.
[16] WENG J S, WENG J, ZHANG J L, et al. DeepChain: auditable and privacy-preserving deep learning with blockchain-based incentive [J]. IEEE Transactions on Dependable and Secure Computing, 2021, 18(5): 2438-2455.
[17] SHARMA P K, PARK J H, CHO K. Blockchain and federated learning-based distributed computing defence framework for sustainable society[J]. Sustainable Cities and Society, 2020, 59: 102220.
[18] ZHAO W J. Blockchain technology: development and prospects[J]. National Science Review, 2019, 6(2): 369-373.
[19] XIE C, KOYEJO S, GUPTA I. Asynchronous federated optimization[EB/OL]. [2022-02-01]. https://arxiv. org/abs/1903. 03934.
[20] JIANG Y H, KONEČNÝ J, RUSH K, et al. Improving federated learning personalization via model agnostic meta learning[EB/OL]. [2022-02-01]. https://arxiv. org/abs/1909. 12488.
[21] LI T, SAHU A K, ZAHEER M, et al. Federated optimization in heterogeneous networks [C]//Proceedings of IEEE Symposium on Machine Learning and Systems. Washington D. C., USA: IEEE Press, 2020: 429-450.
[22] ZHANG M, SAPRA K, FIDLER S, et al. Personalized federated learning with first order model optimization[EB/OL]. [2022-02-01]. https://arxiv. org/abs/2012. 08565.
[23] GAO S, YU T Y, ZHU J M, et al. T-PBFT: an EigenTrust-based practical Byzantine fault tolerance consensus algorithm [J]. China Communications, 2019, 16(12): 111-123.
[24] 孙嘉豪,孟翔斯,张浩运,等. 基于改进PBFT的区块链知识产权保护模型[J]. 计算机工程,2020,46(12):134-141.
SUN J H, MENG X S, ZHANG H Y, et al. Intellectual property protection model using blockchain based on improved PBFT[J]. Computer Engineering, 2020, 46(12): 134-141. (in Chinese)
[25] KIM H, PARK J, BENNIS M, et al. Blockchained on-device federated learning[J]. IEEE Communications Letters, 2020, 24(6): 1279-1283.
[26] 谭敏生,杨杰,丁琳,等. 区块链共识机制综述[J]. 计算机工程,2020,46(12):1-11.
TAN M S, YANG J, DING L, et al. Review of consensus mechanism of blockchain[J]. Computer Engineering, 2020, 46(12): 1-11. (in Chinese)
[27] 王兵,李辉灵,牛新征. 基于综合选举的DPoS共识算法[J]. 计算机工程,2022,48(6):50-56.
WANG B, LI H L, NIU X Z. DPoS consensus algorithm with comprehensive election[J]. Computer Engineering, 2022, 48(6): 50-56. (in Chinese)
[28] NGUYEN T, KIM K. A survey about consensus algorithms used in Blockchain[J]. Journal of Information Processing Systems, 2018, 14(1): 101-128.
[29] MIKAVICA B, KOSTIĆ-LJUBISAVLJEVIĆ A. Blockchain-based solutions for security, privacy, and trust management in vehicular networks: a survey [J]. The Journal of Supercomputing, 2021, 77(9): 9520-9575.

编辑 索书志

(上接第32页)

[38] SCHANZENBACH M, BRAMM G, SCHÜTTE J. reclaimID: secure, self-sovereign identities using name systems and attribute-based encryption[C]//Proceedings of the 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE). Washington D. C., USA: IEEE Press, 2018: 946-957.
[39] DALGARNO B, LEE M J W. What are the learning affordances of 3-D virtual environments?[J]. British Journal of Educational Technology, 2010, 41(1): 10-32.
[40] Ethereum improvement proposals[EB/OL]. [2022-02-03]. https://eips. ethereum. org/erc.
[41] CryptoKitties. CryptoKitties white paper[EB/OL]. [2022-02-03]. https://drive. google. com/file/d/1ikZHnJX yaCH mjmAb5qubaDFREMX1F6PK/view.
[42] Decentraland. Metaverse property[EB/OL]. [2022-02-03]. https://metaverse. properties/buy-indecentraland.
[43] YNAG Q, ZHAO Y T, HUANG H W, et al. Fusing blockchain and AI with Metaverse: a survey[EB/OL]. [2022-02-03]. https://arxiv. org/abs/2201. 03201
[44] Bitcoin for businesses[EB/OL]. [2022-02-03]. https://bitcoin. org/en/bitcoin-for-businesses.
[45] Ethereum. Welcome to Ethereum[EB/OL]. [2022-02-03]. https://ethereum. org/en/.
[46] DAIAN P, GOLDFEDER S, KELL T, et al. Flash Boys 2. 0: frontrunning in decentralized exchanges, miner extractable value, and consensus instability[C]//Proceedings of IEEE Symposium on Security and Privacy. Washington D. C., USA. IEEE Press, 2020: 910-927.
[47] 证券日报. 非同质化权益(NFR)白皮书[EB/OL]. [2022-02-03]. http://www. zqrb. cn/jrjg/hlwjr/2021-10-15/A1634263431812. html.
Securities Daily. Non Homogeneous Equity (NFR) white paper[EB/OL]. [2022-02-03]. http://www. zqrb. cn/jrjg/hlwjr/2021-10-15/A1634263431812. html. (in Chinese)
[48] DUAN H H, LI J Y, FAN S Z, et al. Metaverse for social good: a university campus prototype[C]//Proceedings of the 29th ACM International Conference on Multimedia. New York, USA: ACM Press, 2021: 153-161.

编辑 陆燕菲