文章编号： 1671-7449(2025)01-0046-08

不可靠通信下基于信誉的联邦学习客户端选择

贾惠景[1]，付　芳[2*]，张志才[2]

（1. 山西大学 物理电子工程学院，山西 太原 030051；2. 海南大学 计算机科学与技术学院，海南 海口 570228）

摘　要： 联邦学习作为一种分布式机器学习框架，因其数据隐私保护特性受到广泛关注，然而，恶意客户端和不可靠通信严重影响了其性能与效率。为了解决上述问题，提出了一种不可靠通信下基于信誉的多任务发布者的联邦学习客户端选择机制。首先，使用上行链路模型传输成功概率评估通信可靠性，并考虑了其对聚合模型性能的影响。其次，提出了一种全面的客户端信誉评价方法，并构建了以最大化任务发布者的聚合模型性能以及所选客户端的信誉-价格比为优化目标的客户端选择机制。为了解决该优化问题，将其建模为马尔可夫决策过程，并利用好奇心驱动的深度Q学习算法进行求解。实验结果表明，所提算法显著优于基准算法，对联邦学习的性能有显著改善。

关键词： 联邦学习；不可靠通信；信誉；客户端选择；好奇心驱动的深度Q学习

中图分类号： TP393　　**文献标识码：** A　　**doi:** 10.62756/csjs.1671-7449.2025008

引用格式： 贾惠景，付芳，张志才．不可靠通信下基于信誉的联邦学习客户端选择[J]．测试技术学报，2025，39(1)：46-53.

JIA Huijing，FU Fang，ZHANG Zhicai. Reputation-based client selection for federated learning under unreliable communication[J]. Journal of Test and Measurement Technology，2025，39(1)：46-53.

Reputation-Based Client Selection for Federated Learning Under Unreliable Communication

JIA Huijing[1]，FU Fang[2*]，ZHANG Zhicai[2]

（1. College of Physics and Electronic Engineering，Shanxi University，Taiyuan 030051，China；
2. College of Computer Science and Technology，Hainan University，Haikou 570228，China）

Abstract： Federated learning is a distributed machine learning framework that has received widespread attention for its data protection properties. However，malicious clients and unreliable communication seriously affect its performance and efficiency. To solve the above problems，a reputation-based federated learning client selection mechanism for multi-task publishers in unreliable communication is proposed. Firstly，the communication reliability is evaluated using the uplink model transmission success probability and its impact on the performance of the aggregation model is also considered. Secondly，a comprehensive client reputation evaluation method is proposed and a client selection mechanism with the optimization objective of maximizing the performance of the aggregation model of the task publisher as well as the reputation-price ratio of the selected client is constructed. To solve this optimization problem，it is modeled as a Markov decision process and the curiosity driven deep Q-learning network algorithm is used to achieve optimization. The result shows that the proposed algorithm outperforms the baselines，leading to

收稿日期： 2024-04-22

作者简介： 贾惠景(2000－)，女，硕士生，主要从事联邦学习、强化学习研究。E-mail：jiahuijing2022@163.com。

***通信作者：** 付　芳(1985－)，女，博士，主要从事联邦学习、算网融合研究。E-mail：fufang0621@sxu.edu.cn。

a significant improvement in the performance of federated learning.

Key words: federated learning; unreliable communication; reputation; client selection; curiosity driven deep Q-learning network(CDQN)

0 引 言

近年来，随着物联网技术的发展，移动智能设备和物联网设备数量的急剧增加推动了海量数据的产生，从而加速了人工智能技术的进步[1-2]。传统的机器学习通常需要将原始数据上传至云服务器进行模型训练，存在传输时延大、通信成本高以及用户隐私易泄露等缺点[3]。为了解决上述问题，Google提出了联邦学习(Federated Learning, FL)，这是一种分布式机器学习框架，允许客户端在不暴露其数据的前提下协作训练全局模型，有效降低了通信传输成本，同时保护数据隐私[4]。然而，参与FL训练的客户端并非始终可靠且高质量，可能存在搭便车的低质量客户端[5]，或者有恶意节点[6]实施数据攻击，这些低质量或恶意客户端的参与严重影响了联邦学习性能与效率，尤其在通信资源受限或通信不可靠的情况下，可能会导致模型传输错误甚至模型丢包，进一步影响FL的性能[7]。因此，在不可靠通信环境下，如何合理选择客户端参与FL训练是一个亟待解决的问题。

目前，关于FL客户端选择已提出了多种算法，一些研究侧重于通过具体的选择标准来提高特定性能。例如，选择最早响应的客户端以加快收敛速度[8]，或使用权重差异来评估客户的数据异质性，并优先选择异质性程度较低的客户端[9]。然而，单一选择标准无法全面考虑影响FL性能的多种因素，可能导致选择结果不理想。此外，也有研究围绕优化问题展开。Nishio等[10]构建了一个具有背包约束的最大化问题，并采用贪婪算法在特定截止日期前选择尽可能多的客户端。贺文晨等[11]考虑时延、准确率等多重因素，提出了以准确率最优化为目标的客户端选择模型。Xu等[12]考虑了客户端能量约束下的带宽分配，并利用无线信道信息来实现最优的客户端选择模式。最优化算法通常以准确性、效率、延迟等性能指标为目标，但这往往忽略了恶意节点对联邦学习系统可能造成的负面影响，导致系统在面对恶意行为时表现出较差的鲁棒性和可靠性。为提高系统的鲁棒性，一些研究基于可靠性进行客户端选择。Kang等[13]提出了一种基于多重主观信誉的客户端调度策略，该策略考虑了直接交互和来自其他服务器的评价来选择可信客户端。Zhang等[14]根据模型参数的质量评估客户端可靠性，从客观角度考虑客户端信誉。然而，现有的客户端信誉评价方案通常仅从单一角度进行考虑，可能导致信誉评估结果的不稳定。此外，之前的研究中并未将通信条件的可靠性纳入考量。

因此，本文提出一种在不可靠通信环境下，基于信誉的多任务发布者联邦学习客户端选择机制。首先，基于信道衰落、传输距离等客观信道条件，使用上行链路模型传输成功概率衡量通信的可靠性，并研究了通信条件如何对聚合模型性能产生影响。其次，提出一种更加全面的的信誉评价方法，综合考虑了基于客户端贡献的直接信誉评价和通过PageRank算法[15]整合的来自其他发布者的间接信誉评价。此外，本文以最大化模型性能以及所选客户端的信誉-价格比为目标，将客户端选择问题建模为马尔可夫决策过程(Markov Decision Process, MDP)，并利用好奇心驱动的深度Q学习(Curiosity Driven Deep Q-learning Network, CDQN)算法对优化问题进行求解。最后，基于MNIST数据集，在经典的联邦学习算法FedAvg和FedProx上，对提出的最优化问题模型和CDQN算法进行了仿真实验验证，结果表明，本文算法具有显著优越性。

1 系统架构

1.1 系统模型

如图1所示，本文考虑了一个物联网场景下的FL系统，包括任务发布者层和客户层。

在客户层中，存在一系列具备一定计算和通信资源的FL客户端，它们拥有大量的私有数据，可以被选择参与FL训练，具体客户端集合被表示为$M=\{1,2,\cdots,M\}$。

在任务发布者层中，存在多个任务发布者发布FL任务并选择客户端，具体任务发布者集合被表示为$N=\{1,2,\cdots,N\}$。此外，根据PageRank

算法，将任务发布者之间的交互表示为有向图 $G\langle N,E\rangle$，其中 E 表示任务发布者的指向关系，具体地说，在选择客户端之前，任务发布者会向其他发布者询问建议，若 N_1 指向 N_2 则表示发布者 N_1 向发布者 N_2 请求建议。

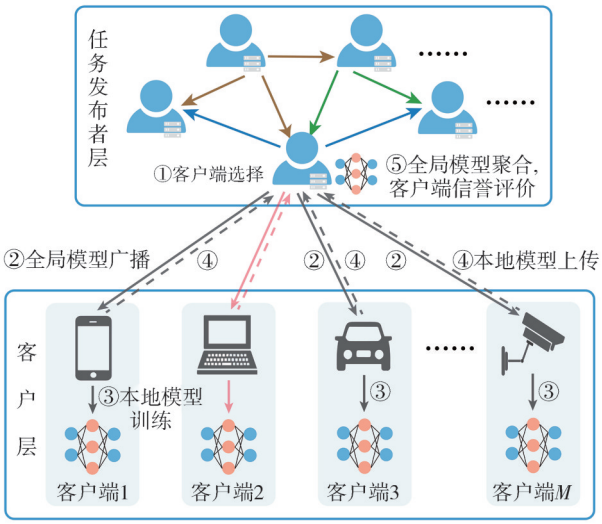


图 1 联邦学习架构

Fig. 1 Federated learning architecture

对于任务发布者 $n\in N$，使用客户端选择指示变量 c_m 衡量客户端 $m\in M$ 是否被选择，若 $c_m=1$ 表示客户端 m 被任务发布者 n 选择参与 FL 训练，$c_m=0$ 表示客户端 m 未被选中。若客户端 m 拥有 D_m 本地数据样本，则 FL 优化目标为寻找最优全局模型 w 以最小化损失函数，具体为

$$\min_{w} F(w)=\sum_{m=1}^{M}\frac{D_m c_m}{\sum_{i=1}^{M}D_i c_i}F_m(w,D_m), \tag{1}$$

其中，客户端 m 的本地损失函数 $F_m(\cdot)$ 定义为

$$F_m(w,D_m)=\frac{1}{D_m}\sum_{x\in D_m}L(w,x)。 \tag{2}$$

1.2 通信模型

在本节中，采用上行链路传输成功概率表征通信可靠性，并假设下行链路参数传输是无差错的[16]，这是因为在下行链路中全局模型是通过广播方式传输，可以占用整个带宽进行通信；而上行链路中客户端拥有更少的通信资源以及更小的发射功率。

首先，用 $h_{n,m}(t)l_{n,m}^{-\frac{\alpha}{2}}$ 表示客户端 m 与任务发布者 n 之间的信道增益，其中 $h_{n,m}(t)\sim\exp(1)$ 表示信道衰落，服从瑞利分布，且在各轮和各设备之间独立同分布，$l_{n,m}$ 为传输距离，α 为路径损耗指数。

那么，客户端 m 在第 t 轮的信噪比（Signal-to-Noise Ratios，SNR）为

$$SNR_m(t)=\frac{P_{n,m}h_{n,m}^2(t)l_{n,m}^{-\alpha}}{B_m(t)N_0}, \tag{3}$$

式中：$P_{n,m}$ 为客户端 m 向任务发布者 n 发送模型的发射功率；$B_m(t)$ 为第 t 轮客户端 m 的带宽；N_0 为噪声功率密度。

设 θ_m 为客户端 m 成功传输的 SNR 阈值，那么客户端 m 在第 t 轮向任务发布者 n 发送模型的成功传输概率为

$$p_{n,m}(t)=Pr\left(\frac{P_{n,m}h_{n,m}^2(t)l_{n,m}^{-\alpha}}{B_m(t)N_0}\geqslant\theta_m\right)=$$
$$Pr\left(h_{n,m}^2\geqslant\frac{\theta_m B_m(t)N_0}{P_{n,m}l_{n,m}^{-\alpha}}\right)=\exp\left(-\frac{\theta_m B_m(t)N_0}{2P_{n,m}l_{n,m}^{-\alpha}}\right)。 \tag{4}$$

对于由不可靠通信产生的传输错误的模型参数，采取全局模型重用策略[17]，即在全局模型聚合过程中使用上一轮的全局模型替代客户端传输错误的本地模型。第 t 轮任务发布者 n 聚合模型描述为

$$w_n(t+1)=\sum_{m=1}^{M}\frac{D_m c_m}{\sum_{i=1}^{M}D_i c_i}\tilde{w}_{n,m}(t), \tag{5}$$

式中：$\tilde{w}_{n,m}$ 为客户端 m 在第 t 轮向任务发布者 n 发送的模型，是一个离散随机变量，服从分布 $\tilde{w}_{n,m}(t)\sim\begin{bmatrix}w_{n,m}(t) & w_n(t-1)\\ p_{n,m}(t) & 1-p_{n,m}(t)\end{bmatrix}$，即 $\tilde{w}_{n,m}(t)$ 以 $p_{n,m}(t)$ 的概率取 $w_{n,m}(t)$，并且以 $1-p_{n,m}(t)$ 的概率取上一轮任务发布者 n 的全局模型 $w_n(t-1)$。

第 t 轮任务发布者 n 聚合模型性能可描述为

$$U_n(t+1)=1-\exp\left(-\varsigma\sum_{m=1}^{M}\tilde{\varphi}_{n,m}(t)p_{n,m}c_m\right), \tag{6}$$

式中：ς 为常数；$\tilde{\varphi}_{n,m}(t)$ 同样是一个离散随机变量，服从分布 $\tilde{\varphi}_{n,m}(t)\sim\begin{bmatrix}\varphi_{n,m}(t) & \varphi_n(t-1)\\ p_{n,m}(t) & 1-p_{n,m}(t)\end{bmatrix}$，并且 $\varphi_{n,m}(t)$ 和 $\varphi_n(t-1)$ 分别为模型 $w_{n,m}(t)$ 以及 $w_n(t-1)$ 的精度。

1.3 信誉模型

由于具有高精度、可靠的训练数据的高信誉客户端在模型训练过程中起着至关重要的作用，

因此，高效准确的声誉计算对于联邦学习中的客户端选择至关重要。为了给予客户端准确的信誉评价，任务发布者将其直接声誉意见与其他发布者的间接声誉意见结合起来生成候选客户端的综合信誉值。

1.3.1　直接信誉

直接信誉的一个重要衡量标准是客户端对全局模型的贡献，这可以反映客户端的模型质量。具体地，使用客户端m本地模型训练后模型精度的增幅来衡量其对全局模型的贡献。因此，客户端m在第t轮对任务发布者n的模型贡献为

$$G_{n,m}(t)=\varphi_{n,m}(t)-\varphi_n(t), \tag{7}$$

式中：$\varphi_n(t)$和$\varphi_{n,m}(t)$分别为第t轮的初始模型精度和客户端m训练后的本地模型精度。需要注意的是，若客户端未被选择，则客户端的模型贡献为0。

为了使客户端的信誉评估具有更高的容错性，只有客户端的模型贡献超过一定阈值时，会引起信誉的变化，具体地，将客户端m在第t轮时的直接信誉定义为

$$R_{n,m}(t)=\begin{cases}R^H(t)+\gamma_1 e^{\frac{G_{n,m}(t)-\beta_1}{\mu_1}}, & G_{n,m}(t)>\beta_1>0,\\ R^H(t)-\gamma_2 e^{\frac{\beta_2-G_{n,m}(t)}{\mu_2}}, & G_{n,m}(t)\leqslant\beta_2<0,\\ R^H(t), & \text{其他},\end{cases} \tag{8}$$

式中：β和β_2为触发直接信誉增加和下降的阈值。使用指数函数$\exp(\cdot)$以及常数$\gamma_1,\gamma_2,\mu_1,\mu_2$来调整贡献对直接信誉的影响强弱，使得客户端在模型性能较差或良好时，都能获得合理的直接信誉评价。

此外，为了使声誉评估更加准确，引入了历史信誉值，并使用时间衰减函数对其进行加权，使历史声誉的参考价值随着时间的推移而降低，故加权历史声誉$R^H(t)$可表示为

$$R^H(t)=\frac{\sum_{l=1}^{l=t-1} e^{-\zeta(t-1)R_{n,m}(t)}}{\sum_{l=1}^{l=t-1} e^{-\zeta(t-1)}}, \tag{9}$$

式中：$e^{-\zeta(t-1)}$为时间衰减函数；ζ为常数。

1.3.2　间接信誉

由于系统中多个任务发布者之间的互动，其他任务发布者对客户端的信誉意见也需要被考虑，同时还需要给予这些信誉评价合适的权重，在设置权重时主要考虑下列因素：

1）偏离平均信誉的程度：如果任务发布者对给定客户的信誉评价与所有发布者提供的平均信誉评价没有太大偏差，则可以认为该任务发布者的评价更趋于真实且更值得信赖，同时也具有越大的参考价值。具体将偏离平均信誉的程度定义为

$$\delta_{n'}(t)=1-\left|R_{n',m}(t)-\frac{\sum_{k\in N_{-n}}R_{k,m}(t)}{N-1}\right|, \tag{10}$$

式中：N_{-n}为除任务发布者n之外的任务发布者集合。当$\delta_{n'}(t)<0$时，说明任务发布者n'的信誉评价不具有参考意义。

2）任务发布者的重要度：由于任务发布者的重要度是存在差异的，应为不同重要度的发布者分配不同的权重，本文使用PageRank算法来计算评价者的重要度，并将其表示为$\sigma_{n'}$。PageRank算法基于以下两个假设：

① 数量假设：如果多个任务发布者指向发布者n，则发布者n的重要度越大；

② 质量假设：重要度大的任务发布者n指向发布者n'，则发布者n'的重要度同样较大。

因此，任务发布者$n'\in N_{-n}$信誉评价的权重为

$$\lambda_{n'}=\frac{\delta_{n'}(t)\sigma_{n'}}{\sum_{k\in N_{-n}}\delta_k(t)\sigma_k}。 \tag{11}$$

客户端m在第t轮时的间接信誉为

$$R_{N_{-n},m}(t)=\sum_{k\in N_{-n}}\lambda_k R_{k,m}(t)。 \tag{12}$$

1.3.3　总信誉

结合直接和间接信誉，第t轮客户端m的总信誉为

$$R_m(t)=\varpi R_{n,m}(t)+(1-\varpi)R_{N_{-n},m}(t), \tag{13}$$

式中：ϖ为常数。

2　基于CDQN的客户端选择

2.1　问题提出

本文优化目标是通过优化客户端选择策略$c(t)=\{c_1(t),c_2(t),\cdots,c_m(t)\}$以使系统模型性能最优，并使客户端信誉值-价格比最大，故优化问题表述为

$$\max_{s(t)}\sum_{t=1}^{T}\left(U_n(t+1)+\sum_{m=1}^{M}\frac{R_m(t)}{B_m(t)}c_m(t)\right), \tag{14}$$

$$s.t.\sum_{t=1}^{T}\sum_{m=1}^{M}c_m(t)B_m(t)\leqslant Budget, \tag{15}$$

$$c_m(t)=\{0,1\}, \tag{16}$$

$$0 \leqslant p_{n,m} \leqslant 1。 \tag{17}$$

约束中，$B_m(t)$为客户端m在第t轮的要价，本文不关注其产生机制，仅将其与数据量相关联。式(15)为预算限制，总支出需小于等于总预算；式(16)中，c_m表示客户端m是否被选择，若$c_m=1$表示客户端m被选择参与FL训练，否则$c_m=0$；式(17)表示模型成功传输概率满足客观条件。

2.2 马尔科夫决策过程

本节将上述客户端选择问题建模为一个MDP，并且将其描述为元组$\langle S, A, r\rangle$，其中具体细节为：

S为状态空间，包含所有客户端的信誉值$R(t-1)=[R_1(t-1),\cdots,R_M(t-1)]$，所有客户端的传输成功概率$p(t-1)=[p_{n,1}(t-1),\cdots,p_{n,M}(t-1)]$以及所有客户端的要价$B(t-1)=[B_1(t-1),\cdots,B_M(t-1)]$。

A为动作空间，包含一个0-1二进制客户端选择策略$c(t)=\{c_1(t),c_2(t),\cdots,c_m(t)\}$。

r为执行某动作后，环境给予的即时奖励，表示为

$$r_{\text{ex}}=U_n(t+1)+\sum_{m=1}^{M}\frac{R_m(t)}{B_m(t)}c_m(t)。 \tag{18}$$

2.3 好奇心驱动的深度Q学习（CDQN）

传统的深度Q学习(Deep Q-learning Network, DQN)通过最大化与环境交互产生的奖励来学习策略[18]，以实现优化目标。在图2所示的CDQN算法中，在DQN的基础上加入了好奇心驱动的内在奖励模块，提高算法的探索能力。

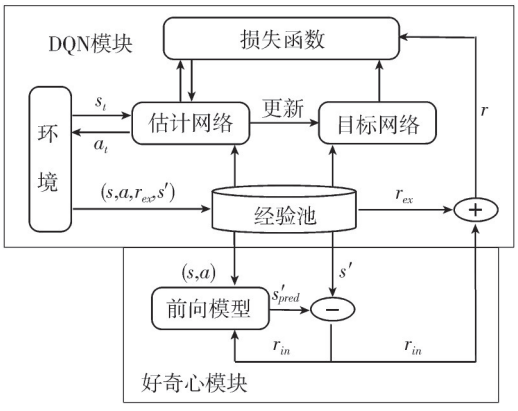


图2 CDQN算法

Fig. 2 CDQN algorithm

2.3.1 好奇心网络

好奇心网络是一个内在奖励生成器，由一个前向模型和一个减法模型组成，分别用于预测状态和生成内在奖励。

前向模型通过从经验回放池提取的样本(s,a,s')预测下一状态，表示为

$$s'_{\text{pred}}=f(a,s;\vartheta), \tag{19}$$

式中：ϑ为前向模型参数，其损失函数为

$$Loss(\vartheta)=\frac{1}{2}\left\| s'-f(a,s;\vartheta)\right\|_2^2。 \tag{20}$$

前向模型参数通过最小化损失函数的策略进行优化，该策略表示为

$$\vartheta \leftarrow \vartheta-\alpha_c(y-f(s,a;\vartheta))\cdot\nabla_\vartheta f(s,a;\vartheta), \tag{21}$$

式中：α_c为好奇心网络中前向模型的学习率，且满足$0<\alpha_c<1$。

此外，好奇心网络生成的内在奖励衡量了智能体对环境的好奇程度，由预测下一状态与实际下一状态的误差得到，表示为

$$r_{\text{in}}=\frac{\tau}{2}\left\| s'-s'_{\text{pred}}\right\|_2^2, \tag{22}$$

式中：$\tau>0$为缩放因子。

2.3.2 DQN网络

DQN的目标是优化状态-动作值函数$Q(s,a)$，输出能够使总奖励最大的最优行动策略，其中总奖励由环境给予的即时奖励和内在奖励共同构成，

$$r=r_{\text{in}}+r_{\text{ex}}。 \tag{23}$$

基于经验回放池提取的样本(s,a,r_{ex},s')，CDQN通过最小化损失函数以更新其网络参数υ，

$$Loss(\upsilon)=E[y-Q(s,a;\upsilon)]^2, \tag{24}$$

式中：E表示数学期望；参数υ更新策略为

$$\upsilon \leftarrow \upsilon-\alpha_q(y-Q(s,a;\upsilon))\cdot\nabla_\upsilon Q(s,a;\upsilon), \tag{25}$$

式中：α_q为DQN网络的参数，且满足$0<\alpha_q<1$。此外，y为目标值，表示为

$$y=\begin{cases} r, & s'\text{为结束状态,}\\ r+\eta\max\limits_{a'\in A}\hat{Q}(s',a';\hat{\upsilon}), & \text{其他。}\end{cases} \tag{26}$$

3 实验结果与分析

3.1 实验设置

1） CDQN算法

回放经验池大小为10^5。CDQN算法的学习率为5×10^4，好奇心网络的学习率为5×10^3。折扣因子以及探索率分别为0.95和0.99。此外，每

隔30轮对目标网络进行更新。

2） 通信环境

客户端个数$M=20$。客户端信噪比阈值和带宽分别在（－20 dBm，20 dBm）和（0.1 MHz，1 MHz）内随机取值，以模拟不同的信道条件，并为客户端产生不同的模型成功传输概率。噪声功率谱密度－174 dBm/Hz。

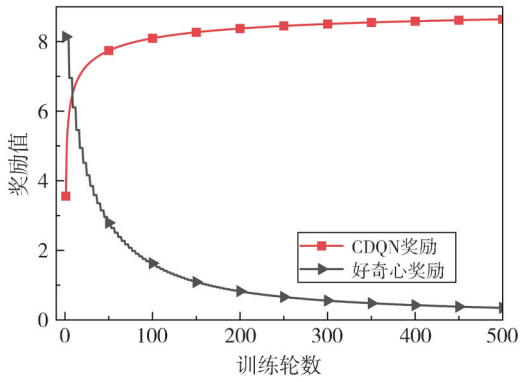
3） 联邦学习

为了评估本文提出的FL客户端选择算法，在MNIST数据集上对经典的FedAvg和FedProx算法进行测试。客户端数据量在（600，6 000）内随机取值。FL训练模型为卷积神经网络，包含2层卷积层、1层最大池化层、2层全连接层以及1层Softmax输出层。在客户端的本地训练中，客户端采用SGD更新，学习率为0.01，本地训练次数为5，训练批大小为20。

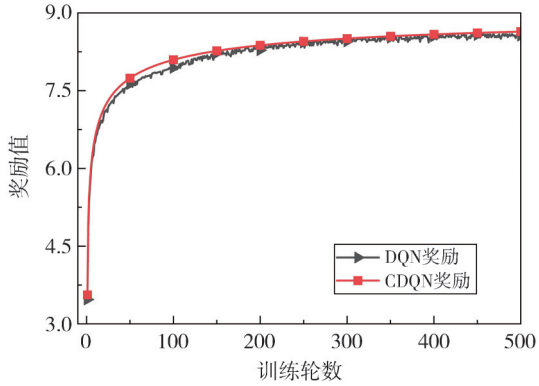
3.2 实验结果

3.2.1 基于CDQN的客户选择算法性能分析

图3（a）展示了CDQN算法的收敛性。由图可知，CDQN奖励值呈现明显的上升趋势并最终趋于稳定，同时由好奇心驱动的内在奖励逐渐降低并接近零。这种现象表明，由内在好奇心奖励和环境提供的外在奖励共同驱动的智能体对环境进行探索后，其预测能力得以提高，导致内在奖励的减少和外在奖励的增加，当环境探索达到充分时，好奇心奖励趋近于零，而外在奖励逐渐趋于稳定。



（a）CDQN收敛性分析



（b）CDQN与DQN对比

图3 算法收敛性对比

Fig. 3 Convergence comparison of algorithms

图3（b）中，对比了本文CDQN算法和传统DQN算法的收敛性，其中CDQN算法显示出更快的收敛速度，同时CDQN收敛后的奖励值略大于DQN，这表明了所提出算法相对于DQN算法的优越性。

3.2.2 客户端选择算法对FL性能的影响

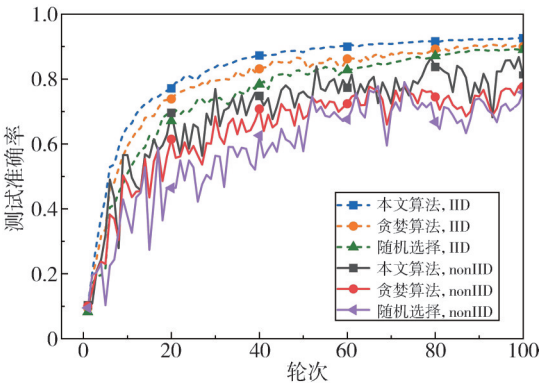
为了验证本文所提客户端选择算法的性能，设置了两个基准实验作为对照：① 随机选择算法，即服务器随机选择客户端参与聚合；②通信贪婪算法，即服务器选择通信条件更好的客户端参与聚合。此外，使用MNIST数据集对FedAvg和FedProx两个经典算法进行实验，且在IID和non-IID两种情况下测试了识别准确率和损失，结果如表1所示。实验结果表明本文的客户端选择算法在学习收敛速度、精度提升和损失降低方面明显优于随机选择算法和通信贪婪算法。

表1 FL实验结果

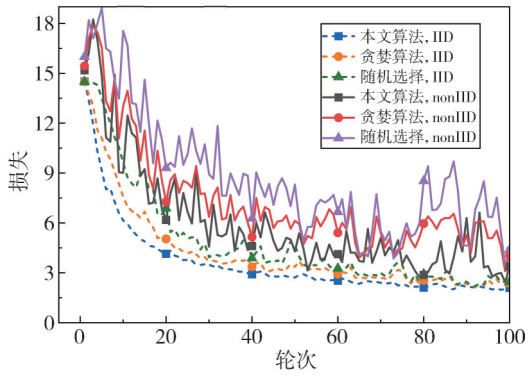
Tab. 1 Experimental results of FL

FL算法	数据设置	测试结果	随机选择	贪婪算法	本文算法
Fed Avg	IID	准确率	0.890 38	0.900 12	0.925 63
		损失	2.455 60	2.356 37	2.096 41
	nonIID	准确率	0.759 50	0.776 63	0.810 75
		损失	4.376 59	4.038 35	3.364 44
Fed Prox	IID	准确率	0.923 38	0.926 05	0. 933 63
		损失	1.860 63	1.749 97	1.635 73
	nonIID	准确率	0.806 50	0.825 00	0.850 75
		损失	3.756 34	3.438 85	2.996 94

图 4 展示了FedAvg算法实验结果。在FedAvg框架下，本文提出的客户端选择算法相较于随机选择算法和通信贪婪算法表现出更优异的性能。具体来说，在IID情况下，本文算法的测试准确率比随机选择算法和通信贪婪算法分别提高了3.959%和2.834%，且损失分别减少了0.359 19和0.259 96；在non-IID分布下，相比较于随机选择算法和通信贪婪算法，本文算法分别实现了6.748%和4.393%的测试准确率提升和1.012 15和0.673 91的损失降低。



（a）准确率



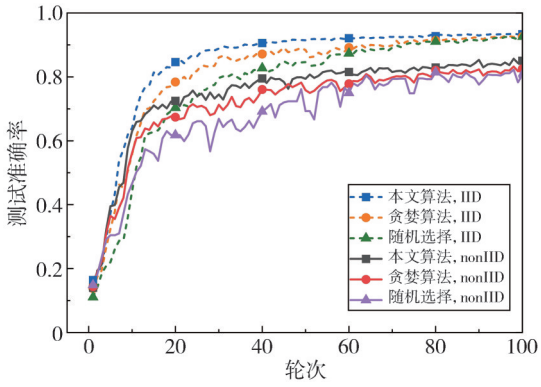
（b）损失

图 4 FedAvg算法实验结果

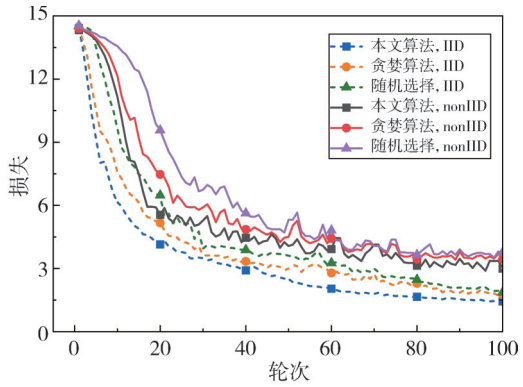
Fig. 4 Experimental results of FedAvg

图 5 展示了FedProx算法实验结果。在FedProx框架下，本文提出的客户端选择算法同样表现出了优势。在IID分布下，本文算法的测试准确率比随机选择算法和通信贪婪算法分别提升了1.110%和0.819%，损失分别降低了0.224 90和0.114 24；在non-IID数据下，测试准确率分别提高了5.487%和3.121%，损失分别降低了0.759 90和0.441 91。

综上所述，本文提出的客户端选择算法在FedAvg和FedProx两种框架下均表现出色，尤其在处理non-IID数据分布时，能够显著提升测试准确率并降低损失。此外，通过对比可以看出，所提出算法在FedAvg框架下对性能的提升作用更明显，最高可提高精度6.748%以及降低损失1.012 15，这是因为相比于FedAvg，FedProx算法本身更具鲁棒性。



（a）准确率



（b）损失

图 5 FedProx算法实验结果

Fig. 5 Experimental results of FedProx

4 结 论

本文提出了一种不可靠通信环境下的联邦学习客户端选择机制，使用模型传输成功概率衡量通信可靠性，并以最大化模型性能以及所选客户端的信誉-价格为目标建立客户端选择体系，此外，将问题建模为MDP，并使用CDQN算法求解。通过模拟客户端的通信信息，并在MNIST数据集上进行测试，证明本文提出的客户端选择算法在学习收敛速度以及模型精度上有较大提升。未来需进一步考虑如何解决联邦学习中通信不可靠以提升联邦学习的性能。

参考文献：

[1] KHAN L U, YAQOOB I, TRAN N H, et al. Edge

computing enabled smart cities：a comprehensive survey[J]. IEEE Internet of Things Journal，2020，7（10）：10200-10232.

[2] LOKSHINA I V，DURKIN B J，LANTING C. The IoT-and big data-driven data analysis services[J]. International Journal of Knowledge Management，2018，14：88-107.

[3] CHEN X，JIAO L，LI W，et al. Efficient multi-user computation offloading for mobile-edge cloud computing[J]. IEEE/ACM Transactions on Networking，2016，24（5）：2795-2808.

[4] MCMAHAN H B，MOORE E，RAMAGE D，et al. Communication efficient learning of deep networks from decentralized data[C]//Proceedings of the 20 th International Conference on Artificial Intelligence and Statistics（AISTATS），2017：1273-1282.

[5] 朱震乾. 联邦学习下的搭便车攻击策略研究[D]. 哈尔滨：哈尔滨工业大学，2022.

[6] 邱晓慧，杨波，赵孟晨，等. 联邦学习安全防御与隐私保护技术研究[J]. 计算机应用研究，2022，39（11）：3220-3231.
QIU Xiaohui，YANG Bo，ZHAO Mengchen，et al. Survey on federated learning security defense and privacy protection technology[J]. Application Research of Computers，2022，39（11）：3220-3231.（in Chinese）

[7] ZANG T，ZHENG C，MA S，et al. A general solution for straggler effect and unreliable communication in federated learning[C]//ICC 2023-IEEE International Conference on Communications，2023：1194-1199.

[8] BUYUKATES B，ULUKUS S. Timely communication in federated learning[C]//IEEE INFOCOM 2021-IEEE Conference on Computer Communications Workshops（INFOCOM WKSHPS），2021：1-6.

[9] ZHANG W，WANG X，ZHOU P，et al. Client selection for federated learning with non-IID data in mobile edge computing[J]. IEEE Access，2021，44（9）：24462-24474.

[10] NISHIO T，YONETANI R. Client selection for federated learning with heterogeneous resources in mobile edge[C]//2019 IEEE International Conference on Communications（ICC），2019：1-7.

[11] 贺文晨，郭少勇，邱雪松，等. 基于DRL的联邦学习节点选择方法[J]. 通信学报，2021，42（6）：62-71.
HE Wenchen，GUO Shaoyong，QIU Xuesong，et al. Node selection method in federated learning based on deep reinforcement learning[J]. Journal on Communications，2021，42（6）：62-71.（in Chinese）

[12] XU J，WANG H. Client selection and bandwidth allocation in wireless federated learning networks：a long-term perspective[J]. IEEE Transactions on Wireless Communications，2021，20（2）：1188-1200.

[13] KANG J，XIONG Z，NIYATO D，et al. Incentive mechanism for reliable federated learning：a joint optimization approach to combining reputation and contract theory[J]. IEEE Internet of Things Journal，2019，6（6）：10700-10714.

[14] ZHANG Q，DING Q，ZHU J，et al. Blockchain empowered reliable federated learning by worker selection：a trustworthy reputation evaluation method[C]//2021 IEEE Wireless Communications and Networking Conference Workshops（WCNCW），2021：1-6.

[15] CHUNG F. A brief survey of PageRank algorithms[J]. IEEE Transactions on Network Science and Engineering，2014，1（1）：38-42.

[16] CHEN M，YANG Z，SAAD W，et al. A joint learning and communications framework for federated learning over wireless networks[J]. IEEE Transactions on Wireless Communications，2021，20（1）：269-283.

[17] YAO J，YANG Z，XU W，et al. GoMORE：global model reuse for resource-constrained wireless federated learning[J]. IEEE Wireless Communications Letters，2023，12（9）：1543-1547.

[18] 张志才，付芳，尹振华. 无人机系统中基于能量效率的资源分配研究[J]. 测试技术学报，2021，35（6）：503-507.
ZHANG Zhicai，FU Fang，YIN Zhenhua. Research on resource allocation based on energy efficiency in UAV system[J]. Journal of Test and Measurement Technology，2021，35（6）：503-507.（in Chinese）