

# 基于演化博弈与中断概率优化的半分布式 联邦边缘学习激励机制

刘雨欣<sup>1</sup>, 王一航<sup>1</sup>, 王思洋<sup>1</sup>, 夏文超<sup>1,2\*</sup>, 赵海涛<sup>1</sup>, 卜宪德<sup>3</sup>

(1. 南京邮电大学教育部泛在网络健康服务系统工程研究中心, 江苏南京 210003; 2. 网络通信与安全紫金山实验室, 江苏南京 211111; 3. 中国电力科学研究院有限公司信息通信研究所, 江苏南京 210003)

**摘 要:** 随着物联网快速发展, 大规模泛在分布终端产生的数据规模激增. 为提升网络智慧服务能力, 采用半分布式联邦边缘学习 (Semi-Decentralized Federated Edge Learning, SD-FEEL) 方式, 通过多个边缘服务器各自协调一个终端簇执行本地更新, 边缘服务器之间周期性共享模型更新, 可以在保障学习性能的同时有效缓解网络拥塞. 然而, 在实际部署中, 缺乏有效激励会导致终端参与学习的积极性不高, 且训练过程中可能发生无线通信中断, 这些因素都会降低模型训练效率. 因此, 本文提出一种面向 SD-FEEL 场景的基于演化博弈与中断概率优化的激励机制. 具体来说, 首先, 设计同时考虑数据大小与质量的终端贡献评估指标, 并据此构建奖励函数激励高质量终端参与训练, 提升全局模型性能并确保激励公平性. 其次, 引入演化博弈框架, 捕捉终端的有限理性和动态决策行为, 均衡边缘服务器负载, 确定种群中关联各边缘服务器的最优比例, 实现种群总效益最大化. 在此基础上, 以最小化无线通信中断概率为目标, 优化终端与边缘服务器的具体关联策略. 仿真结果表明, 所提方法能够有效均衡边缘服务负载, 与随机接入方案和声誉激励机制 (Reputation-Aware Incentive Mechanism, RAIM) 方案相比, 通信中断概率分别降低 32.04% 和 35.55%, 模型准确性分别提升 3.58% 和 4.34%.

**关键词:** 半分布式联邦学习; 激励机制; 演化博弈; 中断概率; 终端贡献评估; 智能物联网

**基金项目:** 江苏省基础研究计划 (自然科学基金) 前沿引领技术基础研究专项 (No. BK20212001); 青年人才托举工程 (No. 2022QNRC001)

**中图分类号:** TP181 **文献标识码:** A **文章编号:** 0372-2112(2025)11-3956-14

**电子学报 URL:** <http://www.ejournal.org.cn> **DOI:** 10.12263/DZXB.20250658

## Incentive Mechanism Based on Evolutionary Game and Interruption Probability Optimization for Semi-Distributed Federated Edge Learning

LIU Yu-xin<sup>1</sup>, WANG Yi-hang<sup>1</sup>, WANG Si-yang<sup>1</sup>, XIA Wen-chao<sup>1,2\*</sup>, ZHAO Hai-tao<sup>1</sup>, BU Xian-de<sup>3</sup>

(1. The Engineering Research Center of Health Service System Based on Ubiquitous Network of Ministry of Education, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210003, China;

2. Purple Mountain Laboratory for Network and Communication Security, Nanjing, Jiangsu 211111, China;

3. Institute of Information and Communication, China Electric Power Research Institute Co., Ltd., Nanjing, Jiangsu 210003, China)

**Abstract:** With the rapid advancement of the Internet of Things, the volume of data generated by large-scale, ubiquitous distributed terminals is surging. To enhance the intelligence of network services, we adopt the semi-decentralized federated edge learning (SD-FEEL) paradigm, in which multiple edge servers each coordinate a cluster of terminals to perform local model updates, while periodically exchanging updates among themselves. This approach preserves learning performance while effectively mitigating network congestion. However, real-world deployments encounter key challenges: inadequate incentives reduce terminals' motivation to participate in training, and wireless communication interruptions during the process can degrade overall training efficiency. To address these issues, this paper proposes an incentive mechanism for SD-FEEL scenarios, leveraging evolutionary game theory and optimization of interruption probabilities. Specifically, first, we

design a terminal contribution metric that incorporates both data quantity and quality, along with a corresponding reward function to encourage participation from high-quality terminals. This not only boosts global model performance but also ensures fairness in incentives. Second, we introduce an evolutionary game framework to model terminals' bounded rationality and dynamic decision-making behaviors. This framework balances edge server loads, determines the optimal proportions of terminals associating with each server within the population, and maximizes the population's overall utility. Building on this foundation, we further optimize specific terminal-to-edge server association strategies with the goal of minimizing the probability of wireless communication interruptions. Simulation results demonstrate that the proposed method can effectively balance the edge service load. Compared to the random access method and the reputation-aware incentive mechanism (RAIM) scheme, the communication interruption probability is reduced by 32.04% and 35.55% respectively, and the model accuracy is improved by 3.58% and 4.34%, respectively.

**Key words:** semi-distributed federated learning; incentive mechanisms; evolutionary games; probability of interruption; terminal contribution evaluation; smart IoT

**Foundation Item(s):** Natural Science Foundation on Frontier Leading Technology Basic Research Project of Jiangsu (No.BK20212001); Young Elite Scientists Sponsorship Program by CAST (No.2022QNR001)

## 1 引言

随着物联网快速发展,网络边缘终端数量激增.据相关预测,至2025年,全球物联网设备数量将增加到750亿台<sup>[1]</sup>,这一增长趋势也导致产生的数据规模日益庞大.机器学习技术可以利用这些海量数据进行建模分析和决策优化<sup>[2]</sup>,赋能智能医疗<sup>[3]</sup>、工业自动化<sup>[4]</sup>和城市管理<sup>[5]</sup>等应用场景.然而传统的集中式机器学习要求终端将采集的数据传输至中心服务器完成模型训练,这一过程面临着诸多挑战<sup>[6]</sup>:一方面,物联网设备采集的数据通常包含个人隐私或商业机密,在传输过程中泄露或超出预期目的的应用,都会引发严重的隐私风险<sup>[7]</sup>;另一方面,将大量的数据集上传至中心服务器进行处理,会加重通信负担<sup>[8]</sup>.

联邦学习<sup>[9]</sup>通过分布式训练设计为解决上述问题提供了可行方案.在联邦学习框架下,终端基于中央服务器分发的全局模型,在本地利用私有数据进行训练,仅将模型参数更新上传,从而在有效保护数据隐私的同时降低通信开销<sup>[10]</sup>.然而,频繁的全局通信会导致网络拥塞和不可避免的延迟.为应对这一问题,文献[11]提出了一种新型的半分布式联邦边缘学习(Semi-Decentralized Federated Edge Learning, SD-FEEL)架构.在该框架下,每个边缘服务器协调一个终端簇执行本地更新与簇内模型聚合,同时,边缘服务器周期性地与相邻的边缘服务器共享其更新模型,以实现簇间的模型聚合.SD-FEEL通过分布式模型聚合和边缘服务器间的低延迟通信,显著提升了训练的效率,降低了网络拥塞发生的可能性.

虽然SD-FEEL在物联网中的应用潜力巨大,但在实际部署中仍面临以下两方面的挑战:一方面,参与训练的终端需消耗大量的计算和通信资源,若缺乏有效的激励,终端可能会缺乏参与的动力,从而影响系统的可持续性.同时,终端在训练过程中,可贡献的数据

数量和数据质量可能不同,这会导致训练效果的差异,进而对全局模型的性能产生影响.此外,边缘服务器将其有限的带宽等资源分配给其下的终端簇,以提升终端上行链路传输效率.但随着簇内终端数量的增加,网络拥塞加剧,导致数据包丢失和重传概率上升,进而增加终端的通信成本.因此,针对这种终端的差异性和相互影响性设计合理的激励机制,优化边缘服务器的终端簇选择方案尤为关键.另一方面,受无线通信系统中路径损耗和阴影效应的影响,终端的通信能力因地理位置差异而异.因此,设计有效的优化策略,实现边缘服务器与终端之间的具体匹配同样重要.

为了激励终端参与模型训练并贡献资源,文献[12~17]设计了多种激励机制.文献[12]针对终端的数据覆盖质量和训练意愿,提出一种二维契约激励机制.通过推导终端的最优努力水平和契约解,解决了逆向选择和道德风险问题.同时采用基于契约的模型聚合方案,提升了模型的泛化性能.文献[13]针对终端感知、计算、传输能力的多维异质性,提出一种基于多维合同的激励机制.利用合同理论的自我揭示特性,解决终端和模型所有者之间信息不对称问题,实现激励的最优匹配.文献[14]针对终端数据的非独立同分布特性,提出多维反向拍卖的模型.依据终端数据的非独立同分布(non-Independent and Identically Distributed, non-IID)程度分组,以最大化社会福利为目标,实现终端选择和激励分配的优化.文献[15]针对终端能量受限问题,提出一种基于在线双边拍卖和最优停止理论的激励机制.通过动态选择能耗最低的终端参与训练,降低能耗和资源浪费.文献[16]针对可靠终端的评估与选择难题,提出结合声誉和贡献度量的反向拍卖的激励机制.任务发布者通过反向拍卖选择高声誉、低出价的终端,并在预算约束下根据终端的实际表现进行激励分配,鼓励高质量终端参与并提高模型质量.文献[17]提出一种基于Stackelberg博

奔的激励机制. 通过动态调整用户设备和基站的定价策略, 优化资源分配和终端的参与度, 从而提升联邦学习的训练精度. 尽管上述研究实现了终端激励, 但它们大多假定终端完全理性并采用一次性决策过程. 然而, 在分布式物联网<sup>[18]</sup>环境中, 终端因全局信息不足或计算能力受限无法实现完全理性. 此外, 终端间效益的相互影响促使决策的动态演化.

演化博弈理论<sup>[19]</sup>通过复制动态(Replicator Dynamics)揭示策略的传播机制, 且假设终端具有有限理性并通过学习和模仿调整策略选择, 适用于分析动态交互环境中个体策略选择和群体行为演化过程. 尽管演化博弈在生物学和经济学领域应用广泛, 但在联邦学习激励机制中的研究较为缺乏. 此外, 现有的基于演化博弈的终端激励的研究大多忽略了无线通信中断问题, 且多局限于分层联邦学习架构上. 事实上, 无线通信中断不仅影响着终端的实际效益, 还会导致边缘服务器无法接收到终端上传的模型参数更新, 进而影响训练的进度和全局模型的精度. 此外, 现有的分层联邦学习架构仍高度依赖与中央服务器的连接, 这往往会导致较高的通信延迟, 从而影响整体训练效率, 且难以扩展到大型网络中.

为此, 本文提出了一种面向SD-FEEL场景的基于演化博弈与中断概率优化的激励机制. 首先, 设计一种同时考虑数据规模和质量终端贡献评估综合指标, 并基于该评估指标建立奖励函数, 激励拥有更多数据样本和更高数据质量的终端参与训练, 提升全局模型性能的同时确保激励的公平性. 然后, 为捕获终端的有限理性和动态决策行为, 引入演化博弈框架. 激励终端参与训练的同时均衡边缘服务器的负载, 确定同一种群内终端关联各边缘服务器的最优比例, 实现种群总体效益的最大化. 通过理论分析证明, 在给定初始条件下, 该博弈模型收敛至唯一均衡解. 最后, 针对无线通信的不稳定性, 引入中断概率作为评估的指标. 基于求解到的种群最优关联比例, 以最小化种群的中断概率为目标, 利用整数线性规划, 优化终端与边缘服务器的具体关联策略, 减小训练过程中的中断轮次, 提升模型准确性. 仿真结果表明, 所提方法能够有效均衡边缘服务器负载, 并且与随机接入方案和基于Stackelberg博弈的声誉激励机制(Reputation-Aware Incentive Mechanism, RAIM)方案<sup>[20,21]</sup>相比, 所提方法使终端平均通信中断概率分别降低32.04%和35.55%, 并使平均模型准确性分别提升3.58%与4.34%.

## 2 系统模型和问题建模

### 2.1 系统模型

考虑的SD-FEEL系统如图1所示, 包含 $J$ 个边缘服

务器, 集合表示为 $\mathcal{J} = \{1, 2, \dots, J\}$ , 以及 $N$ 个终端, 集合表示为 $\mathcal{N} = \{1, 2, \dots, N\}$ . 边缘服务器通过无向连通图 $\mathcal{G} = (\mathcal{J}, \mathcal{E})$ 互连, 其中, 边 $(j_1, j_2) \in \mathcal{E}$ 表示边缘服务器 $j_1$ 和 $j_2$ 通过高速电缆直接连接. 与边缘服务器 $j$ 直接连接的边缘服务器集合定义为 $\mathcal{N}_j = \{i | (i, j) \in \mathcal{E}\}$ , 表示其进行簇间模型共享的范围. 每个终端最多关联至一个边缘服务器, 与边缘服务器 $j$ 关联的终端集合表示为 $\mathcal{C}_j$ , 共同构成一个终端簇. 表1展示了本文的关键参数及其含义.

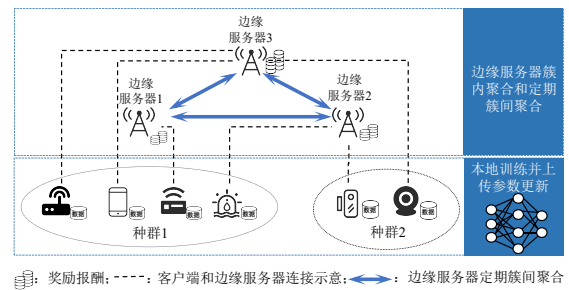


图1 SD-FEEL系统

表1 关键参数表

参数	定义
$T_1, T_2$	边缘服务器的簇内聚合和簇间聚合周期
$D_n^a$	终端 $n$ 的本地数据集
$ D_n^a $	终端 $n$ 的本地数据集大小
$g$	终端 $n$ 的当前本地更新轮次
$n_m$	种群 $m$ 包含的终端数量
$\zeta_j$	由边缘服务器 $j$ 的资源状况所决定的通信成本系数
$r_j$	边缘服务器 $j$ 对种群相对贡献的单位定价
$P_t$	发射信号的功率(单位: dBm)
$P_r$	接收信号的功率(单位: dBm)
$K$	依赖于天线特性和平均信道损耗的常系数
$d$	终端到边缘服务器的距离
$d_0$	天线远场的参考距离
$\gamma$	路径损耗指数

主边缘服务器负责计算终端与边缘服务器的匹配, 由于计算消耗资源, 边缘服务器按照预定顺序轮流担任主边缘服务器. 设边缘服务器的簇内聚合周期 $T_1$ 和簇间聚合周期 $T_2$ 满足 $T_2 = aT_1$ ,  $a$ 为整数. SD-FEEL具体步骤如下:

(1) 初始化. 系统指定一个主边缘服务器及轮换顺序. 模型初始化, 主边缘服务器初始化全局模型, 并广播至其余边缘服务器.

(2) 边缘服务器与终端匹配. 主边缘服务器利用终端本地训练损失值计算数据质量指数, 并基于演化博弈和整数线性规划算法, 计算每个终端与边缘服务器的匹配, 形成终端簇, 并将结果广播至其余边缘服务器. 各边缘服务器与相应终端建立关联, 并将模型参数

下发至簇内终端.

(3)本地训练. 每个终端  $n$  基于最新可用的模型  $w_n$ , 使用随机梯度下降算法和本地数据集  $D_n^a$ , 进行  $T_1$  次本地更新训练. 训练的目标是寻找最优模型参数  $w_n^*$  使损失函数最小<sup>[22]</sup>, 如式(1)所示:

$$w_n^{g*} = \arg \min_{w_n^g} L_n^g \quad (1)$$

其中,  $L_n^g \triangleq \frac{1}{|D_n^a|} \sum_{d \in D_n^a} l(w_n^g, d)$  表示终端  $n$  的损失函数,  $l(w_n^g, d)$  表示数据样本  $d$  在模型参数  $w_n^g$  上的损失函数.

(4)簇内聚合. 每个终端  $n$  完成本地训练后, 将模型参数  $w_n$ 、本地训练损失值和数据量上传至关联的边缘服务器. 边缘服务器  $j \in \mathcal{J}$  收集并存储簇内终端的信息后, 将本地模型聚合形成簇模型  $w_j^G$  后, 下发至簇内终端. 聚合方式如式(2)所示:

$$w_j^G = \sum_{n \in \mathcal{C}_j} \frac{|D_n^a|}{D_j} w_n \quad (2)$$

其中,  $D_j = \sum_{n \in \mathcal{C}_j} |D_n^a|$  为终端簇  $\mathcal{C}_j$  的总数据量.

(5)簇间聚合. 每个边缘服务器  $j$  进行  $a$  次簇内聚合后, 利用共享范围  $\mathcal{N}_j$  内边缘服务器传递的簇模型进行簇间模型聚合, 形成全局模型  $\hat{w}_j^G$ . 聚合方式如式(3)所示:

$$\hat{w}_j^G = \sum_{b \in \mathcal{N}_j \cup \{j\}} \frac{D_b}{\sum_{b \in \mathcal{N}_j \cup \{j\}} D_b} w_b^G \quad (3)$$

同时, 各边缘服务器将簇内终端信息(本地训练损失值、数据量和地理位置等)传播至依据步骤(1)中轮换顺序确定的下一轮主边缘服务器, 用于后续匹配计算.

(6)重复步骤(2)~(5), 直至收敛.

为便于后续分析, 将上述步骤(2)~(5)的组合定义为新一轮全局训练.

## 2.2 数据质量指数模型

考虑到数据质量是影响全局模型训练效果的重要因素之一, 而损失函数值的大小反映了模型训练的效果. 因此, 本文将损失函数值的相对大小作为评估终端数据质量的指标<sup>[23]</sup>. 具体而言, 第  $k$  轮全局训练中, 终端  $n$  的数据质量  $l_n^k$  定义为

$$l_n^k = \frac{\bar{L}_o^k}{L_{o,n}^k} \quad (4)$$

其中,  $\bar{L}_o^k = \frac{1}{N_o} \sum_{n \in \mathcal{N}_o} L_{o,n}^k$  代表所有参与训练的终端的平均损失函数值,  $\mathcal{N}_o \in \mathcal{N}$  代表所有参与训练的终端集合,  $N_o = |\mathcal{N}_o|$  代表所有参与训练的终端数,  $L_{o,n}^k$  是终端  $n$  在第  $k$  轮全局训练中的平均损失函数值, 其定义为

$$L_{o,n}^k = \frac{1}{T_2} \sum_{g=1}^{T_2} L_n^g \quad (5)$$

由式(4)可知,  $l_n^k$  越大, 说明终端  $n$  的数据质量越好. 为了激励终端在每轮训练中持续提供高质量数据, 同时减少数据质量评估的波动, 引入指数衰减系数. 因此, 数据质量的评估不仅与本轮训练的数据质量有关, 还与历史轮次的数据质量有关, 且距离当前轮次越近的轮次影响越大. 定义第  $k$  轮全局训练中终端  $n$  的数据质量指数  $q_n^k$  为

$$q_n^k = \sum_{i=0}^k \mu_i l_n^i \quad (6)$$

其中, 归一化权值  $\mu_i$  为

$$\mu_i = \frac{\mu_i}{\sum_{i=0}^k \mu_i} \quad (7)$$

其中,  $\mu_i = e^{-\lambda(k-i)}$ ,  $\lambda$  是衰减系数, 反映历史数据质量的影响程度,  $l_n^0$  代表初始值, 可以设为终端  $n$  在最近一次训练任务中的平均数据质量.

## 2.3 终端效用模型

为激励终端参与模型训练, 边缘服务器为终端提供一定的奖励. 同时终端选择合适的边缘服务器进行关联, 为其提供模型训练服务. 在第  $k$  轮全局训练前, 主边缘服务器利用各边缘服务器传输的数据计算出所有终端  $n \in \mathcal{N}$  的  $k-1$  轮训练后的数据质量指数  $q_n^{k-1}$ . 利用  $K$ -means 算法, 以数据量大小  $|D_n^a|$  和数据质量指数  $q_n^{k-1}$  作为特征值输入, 将  $N$  个终端划分为  $M$  个种群, 集合表示为  $\mathcal{M} = \{1, 2, \dots, M\}$ . 为消除特征值量纲差异, 在划分前进行归一化处理, 如式(8)所示:

$$z = \frac{y - y_{\min}}{y_{\max} - y_{\min}}, \quad y \in \{|D_n^a|, q_n^{k-1}\} \quad (8)$$

其中,  $y$  为原始特征值,  $y_{\min}$  和  $y_{\max}$  分别为特征值  $y$  的最小值和最大值.

为了方便理论分析, 假设同一种群中的终端拥有相同的数据量  $d_m$ 、数据质量指数  $q_m$  和单位数据所产生的计算成本  $c_m$ . 在  $t$  时刻, 种群  $m$  中选择边缘服务器  $j$  进行关联的终端比例  $x_m^j(t)$  满足  $\sum_{j \in \mathcal{J}} x_m^j(t) = 1$ .

种群  $m$  中每个终端的计算开销为

$$C_m^{\text{cmp}} = T_2 c_m d_m \quad (9)$$

边缘服务器将其通信资源分配给关联的终端, 种群  $m$  中关联边缘服务器  $j$  的每个终端的通信成本计算<sup>[24]</sup>如式(10)所示:

$$C_{m,j}^{\text{com}}(t) = \zeta \left( \sum_{m=1}^M x_m^j(t) n_m \right)^2 \quad (10)$$

因此, 种群  $m$  中关联边缘服务器  $j$  的每个终端产生的总成本为

$$C_m^j(t) = C_m^{\text{cmp}} + C_m^{\text{com}}(t) \quad (11)$$

为了激励终端提供更多、更高质量的数据参与模型训练,边缘服务器根据种群的相对贡献提供奖励.同时,为了补偿终端参与训练的成本并均衡负载,边缘服务器 $j$ 为每个种群 $m$ 提供补偿池 $R_m^j$ .考虑到不同种群拥有的终端数量不同,基于公平性原则,设置 $R_m^j = R_j n_m$ ,其中, $R_j$ 表示边缘服务器 $j$ 提供的单位补偿.因此,边缘服务器 $j$ 给予种群 $m$ 的奖励如式(12)所示:

$$p_m^j(t) = r_j \left[ \alpha \frac{x_m^j(t) D_m}{\frac{1}{M} \sum_{m=1}^M x_m^j(t) D_m} + \beta \frac{x_m^j(t) Q_m}{\frac{1}{M} \sum_{m=1}^M x_m^j(t) Q_m} \right] + R_m^j \quad (12)$$

其中, $D_m = d_m n_m$ , $Q_m = q_m n_m$ 分别表示种群 $m$ 的总数据量和总数据质量指数大小. $\frac{x_m^j(t) D_m}{\frac{1}{M} \sum_{m=1}^M x_m^j(t) D_m}$ , $\frac{x_m^j(t) Q_m}{\frac{1}{M} \sum_{m=1}^M x_m^j(t) Q_m}$

分别反映了种群 $m$ 对边缘服务器 $j$ 数据量和数据质量的相对贡献. $\alpha$ 和 $\beta$ 是加权系数,反映边缘服务器对数据量和数据质量的重视程度,满足 $\alpha + \beta = 1$ .

此外,定义种群 $m$ 中关联边缘服务器 $j$ 的单个终端效用为

$$u_m^j(t) = \frac{p_m^j(t)}{x_m^j(t) n_m} - C_m^j(t) \quad (13)$$

定义种群 $m$ 中所有终端的平均效用为

$$\bar{u}_m(t) = \sum_{j=1}^J x_m^j(t) u_m^j(t) \quad (14)$$

## 2.4 终端通信模型

在无线通信的过程中,终端和边缘服务器之间的通信受到多种信道效应的影响,其中路径损耗和阴影衰落是影响信号强度和传输可靠性的主要因素.路径损耗反映了信号功率随传播距离的衰减,阴影衰落捕捉了由于终端和边缘服务器之间的障碍物所造成的随机衰减.两者共同决定了接收信号功率的强度,接收功率和发射功率之比的分贝值为<sup>[25]</sup>

$$P_{\text{dB}} = P_r - P_t = 10 \log_{10} K - 10\gamma \frac{d}{d_0} - \Psi_{\text{dB}} \quad (15)$$

其中, $K$ 是依赖于天线特性和平均信道损耗的常数, $\Psi_{\text{dB}}$ 是均值为0、方差为 $\sigma_{\Psi_{\text{dB}}}^2$ 的高斯分布随机变量.

从式(15)可以看出,对于任意给定的终端到边缘服务器的距离 $d$ ,接收功率服从对数正态分布.这意味着接收功率 $P_r$ 可能低于最小接收灵敏度 $P_{\text{min}}$ ,从而导致通信链路中断,中断概率表示如式(16)所示:

$$P_{\text{out}} = p(P_r \leq P_{\text{min}}) = 1 - Q \left[ \frac{P_{\text{min}} - (P_t + 10 \log_{10} K - 10\gamma \frac{d}{d_0})}{\sigma_{\Psi_{\text{dB}}}} \right] \quad (16)$$

其中, $Q(z) \triangleq p(Z > z) = \int_z^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{r^2}{2}} dr$ 表示标准正态随机变量 $Z$ 大于 $z$ 的概率.

## 2.5 问题定义

通过上述分析可知,对于种群 $m$ ,其总效用 $U_m$ 不仅由其内部终端与各边缘服务器的关联策略集合所决定,还受到终端关联的边缘服务器 $j \in \mathcal{J}$ 所形成的终端簇 $\mathcal{C}_j$ 内其余终端性能的影响.此外,无线通信过程中的通信中断会导致模型参数上传失败,从而降低总效用.因此,种群中每个终端通过动态调整与各边缘服务器的关联策略,旨在最大化种群总效用,即

$$\begin{aligned} P_0: \max_{\mathbf{S}_m} : U_m(\mathbf{S}_m) &= \sum_{n=1}^{n_m} \sum_{j=1}^J s_{n,j} P_{n,j} u_m^j, \forall m \in \mathcal{M} \\ \text{s.t.} \quad \sum_{j=1}^J s_{n,j} &= 1; \forall n \in \mathcal{N}_m \end{aligned} \quad (17)$$

其中, $\mathbf{S}_m = [s_{1,1}, \dots, s_{1,J}, \dots, s_{n_m,J}]$ 表示种群 $m$ 中终端与边缘服务器的关联策略, $s_{n,j} \in \{0, 1\}$ 表示种群 $m$ 中终端 $n$ 与边缘服务器 $j$ 的关联状态, $s_{n,j} = 1$ 表示进行关联, $s_{n,j} = 0$ 表示不关联, $P_{n,j}$ 表示终端 $n$ 关联边缘服务器 $j$ 后的通信中断概率,与式(16)中的 $P_{\text{out}}$ 一致.

由于上述整数规划问题的求解较为复杂,因此将其分解为 $P_1$ 和 $P_2$ 两个子问题进行近似求解.子问题 $P_1$ 是在不考虑通信中断概率的理想情况下,各种群 $m \in \mathcal{M}$ 依据所有种群与边缘服务器关联的终端比例 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_m, \dots, \mathbf{x}_M]$ ,动态调整自身与各边缘服务器关联的终端比例 $\mathbf{x}_m = \{x_m^1, \dots, x_m^j, \dots, x_m^J\}$ ,以最大化种群总效用,即

$$\begin{aligned} P_1: \max_{\mathbf{x}_m} : U_m(\mathbf{X}) &= \sum_{j=1}^J u_m^j, \forall m \in \mathcal{M} \\ \text{s.t.} \quad \sum_{j=1}^J x_m^j &= 1; \\ 0 \leq x_m^j &\leq 1; \forall j \in \mathcal{J} \end{aligned} \quad (18)$$

在子问题 $P_1$ 解的基础上,子问题 $P_2$ 进一步考虑通信中断概率,使实际效用尽可能地接近最大效用,各种群 $m \in \mathcal{M}$ 在子问题 $P_1$ 的最优比例策略 $x_m^{j*}$ 的约束下,优化其内部终端与边缘服务器的具体关联策略,以最小化自身的中断概率和,实现匹配聚类,即

$$\begin{aligned} P_2: \min_{\mathbf{S}_m} : \sum_{n=1}^{n_m} \sum_{j=1}^J s_{n,j} P_{n,j}, \forall m \in \mathcal{M} \\ \text{s.t.} \quad \sum_{j=1}^J s_{n,j} &= 1; \forall n \in \mathcal{N}_m \\ \sum_{n=1}^{n_m} s_{n,j} &= x_m^{j*} n_m; \forall j \in \mathcal{J} \end{aligned} \quad (19)$$

接下来,首先求解子问题 $P_1$ ,即在不考虑通信中断概率的理想情况下,优化各种群的关联比例策略.

### 3 演化博弈与均衡解

种群  $m$  中终端选择关联不同的边缘服务器  $j \in \mathcal{J}$ , 为其提供模型训练以获取效用  $u_m^j(t)$ . 然而, 由式(10)和式(12)分析可知, 终端的通信成本  $C_{m,j}^{\text{com}}(t)$  和奖励  $p_m^j(t)$  受其所在终端簇  $\mathcal{C}_j$  内其余终端的性能影响. 这意味着终端的效用会随所有终端的策略选择改变而动态变化, 从而驱使终端调整其关联策略. 为了模拟这种相互影响和动态决策过程, 引入演化博弈理论对种群间的交互进行建模. 具体而言, 同一个种群内的终端可以通过比较和交换效用信息, 促使部分终端为提升自身效用而模仿更优策略<sup>[26]</sup>.

#### 3.1 演化博弈均衡分析

为量化动态决策过程, 引入演化博弈理论中的复制动态方程, 描述种群  $m$  选择边缘服务器  $j$  的比例  $x_m^j(t)$  在初始条件  $\mathbf{x}_m(0)$  下, 随时间  $t$  的变化率  $\dot{x}_m^j(t)$ <sup>[27]</sup>. 复制动态方程定义为

$$\dot{x}_m^j(t) = f_m^j(\mathbf{x}_m(t)) = \delta x_m^j(t) (u_m^j(t) - \bar{u}_m(t)), \quad (20)$$

$\forall m \in \mathcal{M}, \forall j \in \mathcal{J}, \forall t$

其中,  $\delta > 0$  表示种群的学习率, 表征种群中的终端模仿并调整其策略的速度. 当选择边缘服务器  $j$  的效用  $u_m^j(t)$  高于种群平均效用  $\bar{u}_m(t)$  时,  $\dot{x}_m^j(t) > 0$ , 选择比例  $x_m^j(t)$  随时间增加; 反之减小. 这反映了种群中具有更高效用的策略会随着时间的推移在种群中变得更加普遍.

为证明在任意初始条件下, 演化博弈均具有唯一均衡解, 提出以下引理.

**引理 1**  $f_m^j(\mathbf{x}_m)$  关于  $x_m^a(t)$  的一阶导数对于  $\forall a \in \mathcal{J}$  均是有界的.

**证明** 为简化说明, 省略时间  $t$  和种群  $m$ , 记  $x_m^j(t)$  为  $x^j$ . 首先, 对  $f_m^j(\mathbf{x}_m)$  求导, 如式(21)所示:

$$\frac{df_m^j(\mathbf{x}_m)}{dx^a} = \delta \left( \frac{dx^j}{dx^a} (u^j - \bar{u}) - x^j \left( \frac{du^j}{dx^a} - \frac{d\bar{u}}{dx^a} \right) \right) \quad (21)$$

其中,

$$\frac{du^j}{dx^a} = \frac{dp^j}{dx^a} - \frac{dx^j}{dx^a} p^j - \frac{dC^j}{dx^a} \quad (22)$$

$$\frac{dC^j}{dx^a} = T_1 c_m D_m \frac{dx^j}{dx^a} + 2\zeta_j n_m \sum_{m=1}^M x^j n_m \quad (23)$$

$$\frac{dp^j}{dx^a} = r_j \alpha M A_1 + r_j \beta M A_2 \quad (24)$$

其中,  $A_1 = \frac{dx^j}{dx^a} D_m - \frac{D_m^2 x^j}{\sum_{m=1}^M x^j D_m - (\sum_{m=1}^M x^j D_m)^2}$ ,

$$A_2 = \frac{dx^j}{dx^a} Q_m - \frac{Q_m^2 x^j}{\sum_{m=1}^M x^j Q_m - (\sum_{m=1}^M x^j Q_m)^2}$$

由于  $x^j \in [0, 1]$  且  $\sum_{m=1}^M x^j \neq 0, r_j, \alpha, M$  等均为有界常数,

因此,  $\frac{dp^j}{dx^a}, \frac{dC^j}{dx^a}$  是有界的, 进而  $\frac{du^j}{dx^a}, \frac{d\bar{u}}{dx^a}$  是有界的. 综上所述,  $\frac{df_m^j(\mathbf{x}_m)}{dx^a}$  是有界的. 证毕.

**定理 1** 对任意的初始条件  $x_m^j(0) \in [0, 1]$ , 满足  $\sum_{j=1}^J x_m^j(0) = 1$ , 复制动态方程(20)存在唯一解.

**证明** 由中值定理可知, 对于任意在  $x_m^1(t), x_m^2(t) \in [0, 1]$ , 存在  $c \in (x_m^1, x_m^2)$ , 使下式成立:

$$\frac{df_m^j(c)}{dx_m^a} = \frac{f_m^j(x_m^1) - f_m^j(x_m^2)}{x_m^1 - x_m^2} \quad (25)$$

由引理 1 可知,  $\frac{df_m^j(c)}{dx_m^a}$  是有界的, 故  $f_m^j(c)$  满足 Lipschitz 条件, 即存在  $\Gamma > 0$ , 使下式成立:

$$|f_m^j(x_m^1) - f_m^j(x_m^2)| \leq \Gamma |x_m^1 - x_m^2| \quad (26)$$

其中,  $\Gamma = \max \left\{ \left| \frac{df_m^j(c)}{dx_m^a} \right| \right\}$ . 根据 Picard 定理可得, 对于任

意的初值  $x_m^j(0) \in [0, 1]$ , 满足  $\sum_{j=1}^J x_m^j(0) = 1$ , 复制动态方程(20)存在唯一解. 证毕.

图 2 直观地展示了复制动态的演化过程. 其中, 橙色区域对应  $u_m^j > \bar{u}_m$  的情况, 由式(20)可得  $\dot{x}_m^j > 0$ , 即种群  $m$  选择边缘服务器  $j$  的比例  $x_m^j$  将在正效用差  $u_m^j - \bar{u}_m$  的驱动下随时间增加. 然而, 由式(13)可知, 随着更多终端接入, 单个终端的效用  $u_m^j$  将减小, 使其向平均效用方向收敛, 导致效用差值  $u_m^j - \bar{u}_m$  逐渐减小,  $x_m^j$  的变化率随之减缓; 蓝色区域的情形则与之相反.

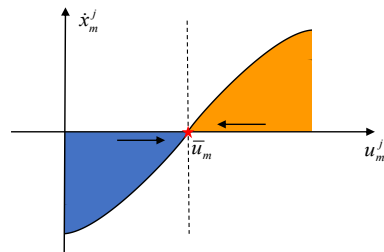


图 2 复制动态演化过程示意图

当所有种群的效用差  $u_m^j - \bar{u}_m = 0$  时, 复制动态丧失进一步演化的动力, 从而达到演化稳定状态, 即演化均衡. 此演化过程适用于任意初始条件, 即动态演化将沿相应路径收敛至该均衡点. 此外, 在任意扰动下, 演化

博弈均会根据此复制动态机制返回该点,即演化均衡解具有唯一性和稳定性.

### 3.2 种群的策略选择算法设计

由定理1可知,演化博弈存在唯一均衡解.主边缘服务器在本地模拟演化博弈过程,依据式(20)中的复制动态方程,推导出种群的策略选择更新过程,寻找种群中终端选择各边缘服务器的最优比例策略,具体如算法1所示.初始时,种群 $m$ 中的终端随机选择某个边缘服务器 $j \in \mathcal{J}$ 进行关联,计算种群选择边缘服务器 $j$ 比例为 $x_m^j(0)$ .在每一时间步 $t$ ,边缘服务器收集簇内终端性能依据式(13)计算每个终端的效用 $u_m^j(t)$ ,并传播至终端.终端通过种群内效用信息的交换,由式(14)计算种群的平均效用 $\bar{u}_m(t)$ .根据复制动态方程式(20),更新种群 $m$ 中选择边缘服务器的比例 $x_m^j(t)$ ,从而实现策略的动态演化,并得到种群中终端选择各边缘服务器比例的均衡解.

#### 算法1 策略选择更新算法

输入:种群和边缘服务器的特征参数,最大迭代次数 $T_{\max}$

输出:种群 $m$ 选择各边缘服务器的终端向量 $\mathbf{n}_m$

1. 初始化: $t \leftarrow 0$ ,种群 $m$ 中终端随机连接一个边缘服务器 $j$
2. WHILE  $t < T_{\max}$  DO
3. FOR  $m \in \mathcal{M}$  DO
4. FOR  $j \in \mathcal{J}$  DO
5.  $u_m^j \leftarrow p_m^j(x_m^j, n_m) - C_m^j$
6. END FOR
7. 利用式(14)计算种群 $m$ 的平均效用 $\bar{u}_m$
8. FOR  $j \in \mathcal{J}$  DO
9.  $x_m^j \leftarrow x_m^j + \delta x_m^j (u_m^j - \bar{u}_m)$
10. END FOR
11.  $x_m^j \leftarrow x_m^j / \sum_{j=1}^J x_m^j$
12. END FOR
13.  $t \leftarrow t + 1$
14. END WHILE

接下来求解问题 $P_2$ .

## 4 种群通信中断概率优化

在SD-FEEL中,种群划分基于终端的性能评估,即数据量和数据质量指数,但终端在地理位置上是分散且随机的.因此,种群 $m \in \mathcal{M}$ 中每个终端到边缘服务器 $j \in \mathcal{J}$ 的距离不同,导致通信中断概率存在差异.通信中断会导致边缘服务器无法接收到终端上传的模型参数更新,进而影响训练的进度和全局模型的精度,同时种群 $m$ 获得该轮的奖励 $p_m^j$ 也因此减小.为此,本文提出一种优化模型,在演化博弈均衡解的限制下,主边缘服务器通过优化种群 $m$ 中终端 $n$ 和边缘服务器 $j$ 的关

联策略,以最小化种群中断概率和,实现基于匹配的聚类.

### 4.1 种群策略选择分配

演化博弈结果提供了种群 $m$ 选择边缘服务器 $j$ 的比例 $x_m^j$ ,然而,实际中终端的数量为整数,因此,需要将种群 $m$ 的总终端数 $n_m$ 分配到 $J$ 个边缘服务器中.设种群 $m$ 中选择边缘服务器 $j$ 的终端数量 $N_{m,j}$ ,寻找最优的值使结果的百分比 $N_{m,j}/n_m$ 与演化博弈均衡解 $x_m^j$ 偏差最小. $N_{m,j}$ 满足以下条件:

$$\begin{aligned} \arg \min_{\mathbf{n}_m} \sum_{j=1}^J \left| \frac{N_{m,j}}{n_m} - x_m^j \right|, \forall m \in \mathcal{M} \\ \text{s.t. } \sum_{j=1}^J N_{m,j} = n_m \end{aligned} \quad (27)$$

为得到种群的最优分配结果,设计递归搜寻算法求解最优的 $\mathbf{n}_m = [N_{m,1}, N_{m,2}, \dots, N_{m,J}]$ .具体流程如算法2所示.通过递归枚举种群 $m$ 分配到每个边缘服务器的终端数量,计算对应的结果偏差,并选择偏差最小的分配方案.

### 4.2 中断概率优化

在确定了种群 $m$ 中选择边缘服务器 $j$ 的终端数量 $N_{m,j}$ 后,进一步优化种群中终端与边缘服务器的关联策略,以最小化种群中断概率和.终端 $n$ 关联边缘服务器 $j$ 后的中断概率为 $P_{n,j}$ ,由式(16)计算得出,即

$$P_{n,j} = P_{\text{out}}(d_{n,j}) \quad (28)$$

其中, $d_{n,j}$ 是终端 $n$ 到边缘服务器 $j$ 的欧式距离,计算公式为

$$d_{n,j} = \sqrt{(x_n - x_j)^2 + (y_n - y_j)^2} \quad (29)$$

其中, $(x_n, y_n)$ 和 $(x_j, y_j)$ 分别表示终端 $n$ 和边缘服务器 $j$ 的位置坐标.

将上述的计算和定义带入 $P_2$ 中,得到

$$\begin{aligned} \arg \min_{\mathbf{s}_m} \sum_{n=1}^{n_m} \sum_{j=1}^J s_{n,j} P_{n,j}, \forall m \in \mathcal{M} \\ \text{s.t. } \sum_{j=1}^J s_{n,j} = 1; \forall n \in \mathcal{N}_m \\ \sum_{n=1}^{n_m} s_{n,j} = N_{m,j}; \forall j \in \mathcal{J} \end{aligned} \quad (30)$$

为解决式(30)中的优化问题,设计基于整数线性规划的中断概率优化算法.具体流程如算法3所示.该算法首先计算每个终端与各边缘服务器的欧式距离和中断概率,并定义二进制决策变量 $s_{n,j} \in \{0, 1\}$ .随后,对每个种群 $m$ 构建目标函数和约束条件.最后,利用整数线性规划模型求解终端-边缘服务器最优的关联策略.

算法复杂度分析如下:第一步的计算总复杂度为

**算法 2 递归搜索算法**

输入: 种群集合  $\mathcal{M}$ , 边缘服务器的数量  $J$ , 种群  $m$  总终端数集合

$\{n_m | m \in \mathcal{M}\}$ , 种群  $m$  选择边缘服务器  $j$  的比例集合  $x_m =$

$\{x'_m | m \in \mathcal{M}, j \in \mathcal{J}\}$

输出: 种群  $m$  选择各边缘服务器的终端数向量  $n_m$

```

1. 初始化: 暂存矩阵 number_j 和最优数量矩阵 best_number
2. FOR  $m \in \mathcal{M}$  DO
3. 初始化最小偏差 min_deviation  $\leftarrow \infty$ 
4. 调用递归函数 number( $m, 0, n_m, x_m, J, \text{number\_j}, \text{best\_number},$ 
min_deviation)
5. END FOR
6. number( $m, j, n_m, x_m, J, \text{number\_j}, \text{best\_number}, \text{min\_deviation}$ )//递归
函数
7. IF  $j = J - 1$  THEN
8. number_j( $j$ )  $\leftarrow n_m - \sum_{j=0}^{j-2} \text{number\_j}(j)$ 
9. 计算偏差 deviation  $\leftarrow \sum_{j=0}^{j-1} \text{abs}\left(\frac{\text{number\_j}(j)}{n_m} - x_m(j)\right)$ 
10. IF deviation < min_deviation THEN
11. min_deviation  $\leftarrow$  deviation
12. best_number( $m$ )  $\leftarrow$  number_j
13. END IF
14. ELSE
15. min_val  $\leftarrow$  min{0, int( $x_m(j)n_m - 2$ )}
16. max_val  $\leftarrow$  max{int( $x_m(j)n_m + 2$ ),  $n_m$ }
17. FOR val  $\leftarrow$  min_val TO max_val DO
18. number_j( $j$ )  $\leftarrow$  val
19. number( $m, j + 1, n_m, x_m, J, \text{number\_j}, \text{best\_number},$ 
min_deviation)
20. END FOR
21. END IF

```

$O\left(\sum_{m \in \mathcal{M}} n_m \times J\right) = O(N \times J)$ , 其中  $N = \sum_{m \in \mathcal{M}} n_m$  为终端总数量;

第二步的总复杂度为  $O\left(\sum_{m \in \mathcal{M}} n_m \times J\right) = O(N \times J)$ ; 第三步将

问题表述为一个 0-1 型整数线性优化问题, 对于种群  $m$  而言, 决策变量总数  $r = n_m \times J$ , 约束条件数  $v = 2(n_m \times J)$ . 通常, 采用精确算法求解时, 在最坏情况下的时间复杂度呈

指数型  $O\left(2^{\left(1 - \text{poly}\left(\frac{r}{v}\right)\right)r}\right)$ <sup>[28]</sup>. 然而, 在实际物联网场景中, 可

采用启发式算法简化运算, 实现多项式型时间复杂度, 从而适用于大规模终端部署<sup>[29]</sup>. 例如, 基于遗传算法的求解时间复杂度近似为  $O(r^{1.025})$ <sup>[30]</sup>, 因此第三步的总时间复杂度

近似为  $O\left(\sum_{m \in \mathcal{M}} (n_m \times J)^{1.025}\right)$ , 进而算法 3 的总时间复杂度

近似为  $O\left(\sum_{m \in \mathcal{M}} (n_m \times J)^{1.025} + N \times J\right)$ . 该复杂度随种群  $m$  的终端数量  $n_m$  呈 1.025 次幂增长, 随总终端数  $N$  呈线性增长, 不会因终端规模扩大而导致指数级爆炸, 从而确保在实际部署中保持计算开销可控.

**算法 3 中断概率优化算法**

输入: 种群集合  $\mathcal{M}$ , 边缘服务器的集合  $\mathcal{J}$ , 种群  $m$  总终端数集合

$\{n_m | m \in \mathcal{M}\}$ , 种群  $m$  中的终端集合  $\{\mathcal{N}_m | m \in \mathcal{M}\}$ , 种群  $m$  中选择各边缘服务器的终端数  $n_m$

输出: 优化后的策略选择状态  $\mathcal{S}$

```

1. 初始化问题实例为最小化问题
2. FOR  $m \in \mathcal{M}$  DO
3. FOR  $n \in \mathcal{N}_m$  DO
4. FOR  $j \in \mathcal{J}$  DO
5. 计算终端到该边缘服务器的欧式距离  $d_{n,j}$ 
6. 根据式(28)计算中断概率  $P_{n,j}$ 
7. 定义决策变量  $s_{n,j}$  为二进制变量
8. END FOR
9. END FOR
10. 添加目标函数:  $\sum_{n=1}^{n_m} \sum_{j=1}^J s_{n,j} P_{n,j}$ 
11. FOR  $n \in \mathcal{N}_m$  DO
12. 添加约束条件 1:  $\sum_{j=1}^J s_{n,j} = 1$ //终端只能关联一个边缘服务器
13. FOR  $j \in \mathcal{J}$  DO
14. 添加约束条件 2:  $\sum_{n=1}^{n_m} s_{n,j} = N_{m,j}$ 
15. 求解整数线性规划问题
16. END FOR

```

## 5 实验结果与分析

### 5.1 实验设置

考虑的 SD-FEEL 系统包含 3 个边缘服务器和 90 个终端, 随机部署于  $500 \times 500 \text{ m}^2$  的区域内, 且终端与边缘服务器的距离满足远场条件  $d_{n,j} > 10 \text{ m}$ . 实验参数设置参考文献[22, 25, 31], 边缘服务器簇内聚合周期  $T_1 = 5$ , 簇间聚合周期  $T_2 = 5$ . 终端的发射功率  $P_t = 10 \text{ dBm}$ , 最小接收灵敏度  $P_{\text{min}} = -110 \text{ dBm}$ , 阴影衰落服从标准差为  $\sigma_{\text{sh}}^2 = 3.65 \text{ dB}$  的对数正态分布, 路径损耗指数  $\gamma = 3.71$ . 实验基于 CIFAR-10<sup>[32]</sup>、MNIST<sup>[33]</sup> 以及 Tiny-ImageNet<sup>[34]</sup> 数据集. 其中, CIFAR-10 与 MNIST 数据集采用包含 2 层卷积、2 层池化和 3 层全连接的卷积神经网络进行训练; Tiny-ImageNet 数据集则采用 ResNet-18 网络进行训练. 考虑终端数据满足 IID 和 non-IID 2 种场景: IID 场景下, 数据样本随机分配至 90 个终端, 其中 CIFAR-10 和 MNIST 数据集的终端数据量服从 [400~760] 均匀分布, Tiny-ImageNet 数据集的终端数据量服从 [755~1 467] 均

匀分布;non-IID场景下,数据按类别依据Dirichlet分布不均匀分配给终端,其中分布参数 $\alpha=0.5$ .其余实验参数如表2所示.

表2 实验参数配置表

参数	值	参数说明
$N$	90	终端的数量
$M$	5	种群数量
$J$	3	边缘服务器数量
$\alpha, \beta$	0.5	数据量和数据质量的加权系数
$c_m$	0.001	种群 $m$ 中单位数据所产生的计算成本
$\zeta_j$	[0.01,0.03]	边缘服务器 $j$ 的通信成本系数
$R_j$	[3,5]	边缘服务器 $j$ 为每个种群提供的单位补偿支付
$r_j$	[200,400]	边缘服务器 $j$ 在对种群相对贡献的单位定价

## 5.2 演化博弈仿真结果

### 5.2.1 种群划分

为获取终端 $n$ 在最近一次历史训练任务中的平均数据质量 $l_n^0$ ,在全局训练开始之前,所有终端进行一次本地训练,返回数据质量值.利用K-means算法的聚类结果如图3所示.其中同一颜色表示同一种群,红色的叉表示聚类中心.结果表明,具有相似的数据量和数据质量指数的终端被聚合成一个种群.

### 5.2.2 演化博弈的稳定性

为分析演化博弈的稳定性,设置初始条件使每个群体的终端均匀分配到3个边缘服务器中,即初始选择比例 $x_m^j(0)=1/3, \forall m \in \mathcal{M}, \forall j \in \mathcal{J}$ .图4展示了种群1~5的边缘服务器选择比例 $x_m^j$ 随时间的变化.以种群1为例,初始阶段,因效用激励,终端选择边缘服务器3的比例快速

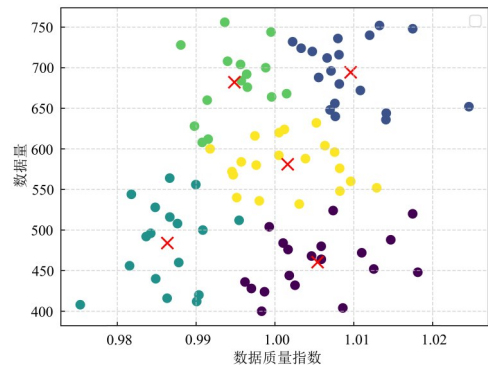


图3 种群聚类结果图

上升,但随后因拥塞效应和奖励池被瓜分的约束,最终稳定在了0.41左右,表明策略选择达到均衡.其他种群的演化趋势类似,验证了式(20)中复制动态方程的稳定性.

图5显示了种群 $m$ 中终端选择边缘服务器 $j$ 所获得的效用随时间的变化,其中实线、虚线、点线分别对应选择边缘服务器1、2、3所获得的效用.边缘服务器3的奖励最高,因此初始阶段选择边缘服务器3的终端效用较高.然而,随着其余终端的模仿、学习并选择加入边缘服务器3,奖励池被瓜分,拥塞导致通信成本增加,从而效用下降.约经150个时间步后,选择各边缘服务器的效用趋于一致,终端不再具有改变策略选择的动力,系统达到稳定状态.

为分析种群学习率对策略演化的影响及其敏感性,本文在相同初始条件下,即每个种群选择每个边缘服务器的比例均为1/3,通过改变种群的学习率 $\delta$ ,研究策略选择的动态演化过程.图6展示了种群1选择边缘

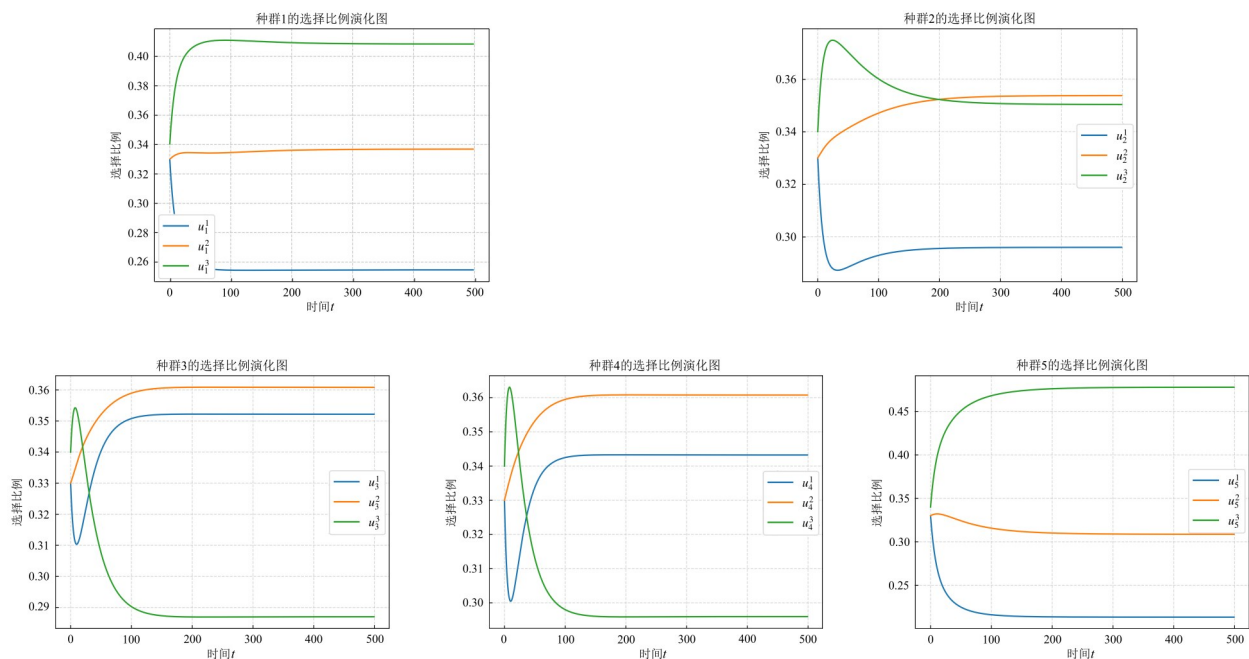


图4 种群的选择比例演化图

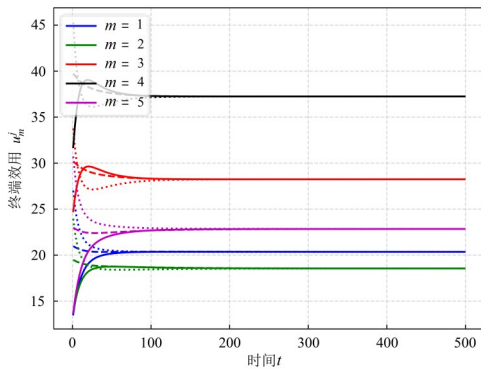


图5 种群平均效用演化图

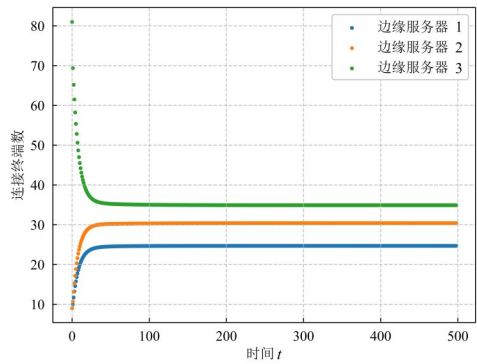


图7 选择各边缘服务器的终端数随时间的变化图

服务器 1 的比例  $x_1^i$  随时间步的变化曲线. 实验结果表明, 学习率控制着策略适应的速度, 从而影响演化均衡的收敛速度. 在一定范围内, 随着学习率的提高, 系统收敛速度加快, 从而更快逼近演化均衡状态, 且不改变均衡解; 然而, 当学习率超出该范围, 系统则会出现振荡的现象, 影响收敛的过程. 这与复制动态方程的数值特性一致, 学习率相当于迭代的步长, 调控着逼近效率而非均衡点本身, 但若迭代步长过大, 则会导致迭代振荡并阻碍稳定收敛. 因此, 选择适当的学习率至关重要, 以平衡收敛速度与稳定性.

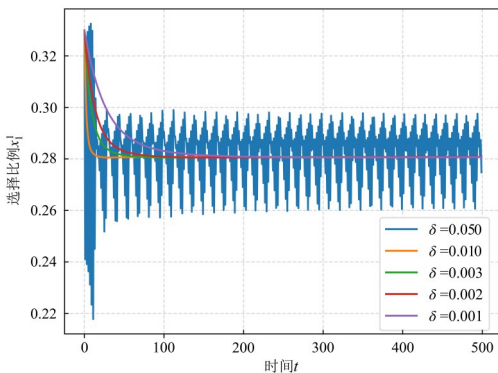


图6 不同学习率条件下种群选择比例演化图

种群 3、4、5 中没有终端选择边缘服务器 1. 图 8 展示了不同初始状态下, 种群 1 和种群 2 选择边缘服务器 1 的比例的演化轨迹, 如初始状态 1 中, 种群 1 和种群 2 初始选择边缘服务器 1 的比例分别为 0.1 和 0.2. 结果表明, 在不同的初始状态下, 种群选择比例的演化过程轨迹均沿着箭头方向收敛至平衡点, 验证了演化博弈的唯一性, 与定理 1 分析一致.

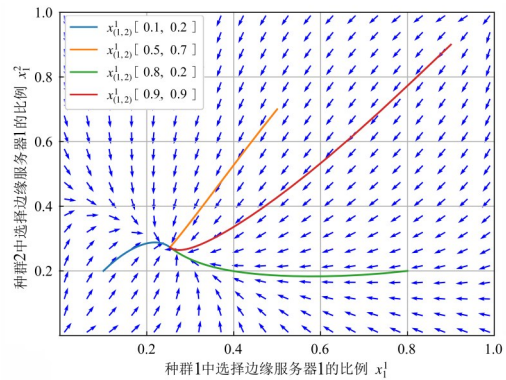


图8 复制子动力学相位图

### 5.2.3 演化博弈对负载的均衡效果

为进一步分析演化博弈对边缘服务器负载的均衡效果, 设置了一个极端的初始条件: 90% 的终端都因高奖励的吸引, 选择边缘服务器 3. 图 7 展示了选择各边缘服务器的终端数随时间的变化. 初始条件下, 边缘服务器 3 的负载高达 81 个终端, 拥塞严重. 随着演化博弈的进行, 终端逐渐转向边缘服务器 1 和边缘服务器 2, 最后 3 个边缘服务器的负载分别为 24、30、36, 负载的标准差从初始的 36.062 降至 4.49, 实现了负载的均衡.

### 5.2.4 演化博弈的唯一性

为验证演化博弈解的唯一性, 通过调整种群初始选择边缘服务器的比例进行实验. 为便于分析, 假设种

### 5.3 中断优化仿真结果

为评估中断优化模型的性能, 将所提优化方案与以下 3 种方案进行对比:

(1) 理想方案: 假设终端与边缘服务器之间的无线通信链路处于理想状态, 通信中断概率为零, 所有终端均能成功上传更新后的模型参数.

(2) 随机连接方案: 除不执行中断概率优化算法外, 其余设置与所提优化方案一致. 在确定各种群选择每个边缘服务器的终端数量后, 采用随机分配的方式进行连接.

(3) RAIM 方案: 利用基于 Stackelberg 博弈的声誉感知激励机制进行终端与边缘服务器的匹配.

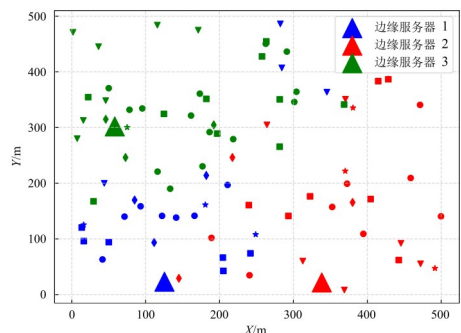
表 3 展示了 5 次实验中 5 个种群分别采用所提优化方案和随机连接方案的平均中断概率对比. 实验结果表明, 相同的实验条件下, 所提优化方案的平均中断概

率显著低于随机连接方案,如种群1平均中断概率从53.5%降至23.7%.

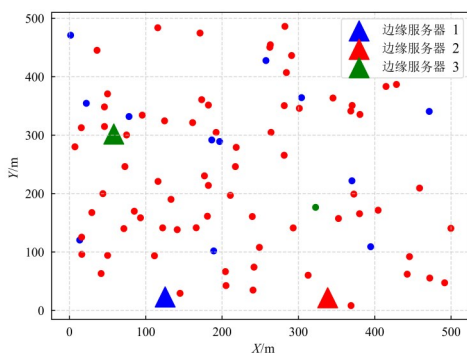
表3 种群平均中断率对比表 单位:%

平均中断率	所提优化方案	随机连接方案
种群1	23.7	53.5
种群2	21.1	53.7
种群3	26.0	58.6
种群4	22.9	56.3
种群5	23.1	55.1

图9展示了所提优化方案与RAIM方案中终端与边缘服务器的连接状态. 其中,同一颜色表示终端连接至同一边缘服务器;在所提优化方案中,同一图标进一步表示属于同一种群. 由图可见,相较于RAIM方案,所提优化方案能够实现更均衡的边缘服务器负载分配,同时显著缩短终端与所连边缘服务器的距离,从而有效降低通信中断概率. 实验结果显示,与RAIM方案相比,所提优化方案将终端的平均中断概率从58.95%降低至23.4%.



(a) 所提优化方案

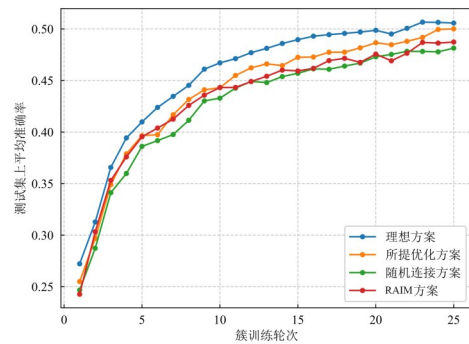


(b) RAIM方案

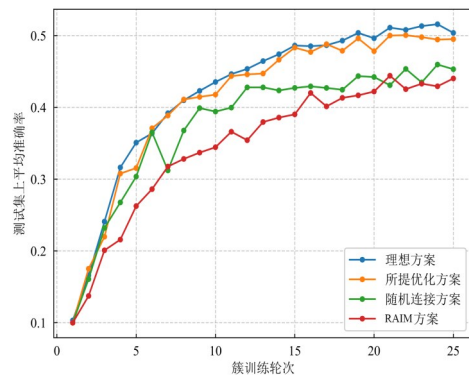
图9 边缘服务器连接图

图10、图11和图12分别展现了理想方案、所提优化方案、随机连接方案以及RAIM方案在不同数据集以及数据分布上的测试准确率曲线. 表4进一步列出这

些方案在全局训练结束后,在不同数据集的测试集上的准确率对比结果. 实验结果表明,所提优化方案的最终准确率高于随机连接方案和RAIM方案,更接近于理想方案. 因为优化后具有更低的通信中断率,在每轮簇内聚合时,有更多的终端上传模型参数更新,聚合后的模型性能更好.



(a) CIFAR-10

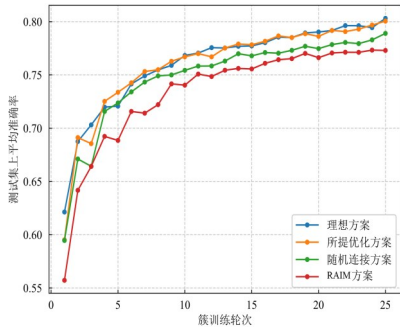


(b) CIFAR-10 non-IID

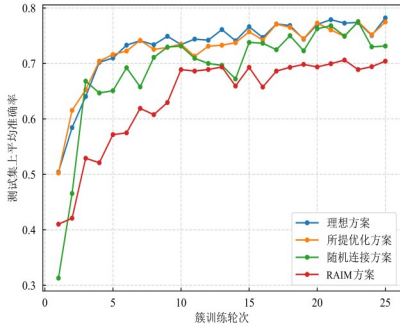
图10 模型在CIFAR-10 IID和CIFAR-10 non-IID测试集上准确率

### 5.4 模型综合评估

基于上述仿真结果表明,所提方案通过演化博弈与中断概率优化机制的协同作用,在实际部署中展现出显著的工程应用潜力. 演化博弈算法能够有效均衡边缘服务器负载,避免单点故障并提升系统的稳定性. 同时,中断优化算法显著降低了通信中断率,与随机接入方案和RAIM方案相比,所提优化方案将平均中断概率分别降低32.08%和35.55%,从而保障更多终端成功上传参数更新,提升簇内聚合的质量,并提高资源的利用率与模型的准确率. 尽管准确率提升幅度相对有限,但在大规模样本的累积效应下,该提升在SD-FEEL部署中仍具有重要的工程意义. 如在边缘医疗检测等经典应用场景中,所提方案可有效减少误判事件的发生,从而提高系统的可靠性和物联网(Internet of Thing, IoT)边缘设备的整体性能.

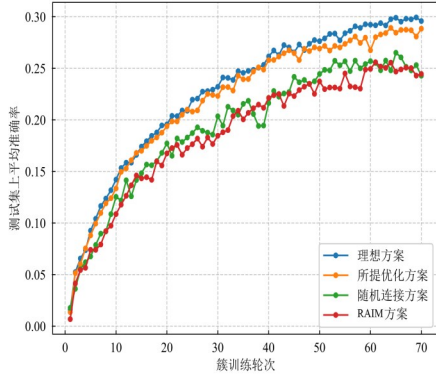


(a) MNIST IID

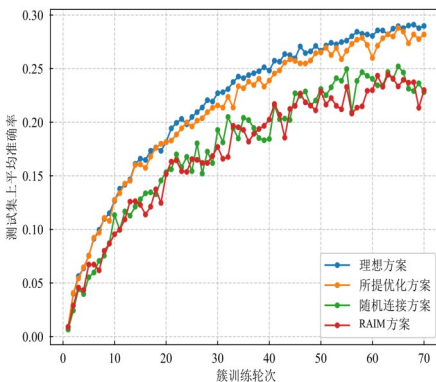


(b) MNIST non-IID

图 11 模型在 MNIST IID 和 MNIST non-IID 测试集上准确率



(a) Tiny-ImageNet IID



(b) Tiny-ImageNet non-IID

图 12 模型在 Tiny-ImageNet IID 和 Tiny-ImageNet non-IID 测试集上准确率

表 4 不同方案在全局训练结束后的准确率对比表 单位: %

模型	CIFAR-10		MNIST		Tiny-ImageNet	
	IID	non-IID	IID	non-IID	IID	non-IID
理想方案	50.57	50.38	80.31	78.21	29.58	28.98
所提优化方案	50.01	49.51	80.05	77.48	28.84	28.19
随机连接方案	48.14	45.32	78.9	73.14	24.28	22.81
RAIM方案	48.73	44.03	77.3	70.42	24.45	23.00

## 6 结论

为实现无线网络下高效的、可持续性的SD-FEEL, 本文提出了一种基于演化博弈与中断概率优化的激励机制. 首先, 基于演化博弈的复制动态方程, 模拟了终端间的交互和策略选择的演化过程, 并从理论上证明了均衡解的唯一性. 然后, 引入中断优化模型通过整数线性规划最小化中断概率和, 确定了终端与边缘服务器的具体关联策略. 实验结果验证了演化博弈均衡解的唯一性和稳定性以及对边缘服务器负载均衡的有效性, 与随机连接方案相比该机制显著降低了平均中断概率并提升了模型性能. 在未来工作中, 进一步从多层次博弈的角度出发, 设计更全面、更有效的激励机制.

## 参考文献

- [1] SCHILLER E, AIDOO A, FUHRER J, et al. Landscape of IoT security[J]. Computer Science Review, 2022, 44(1): 1-18.
- [2] HUSEN A, CHAUDARY M H, AHMAD F, et al. Holistic learning and optimization scheme for wireless networks[J]. Transactions on Emerging Telecommunications Technologies, 2024, 35(10): e5036.
- [3] DUTTA S. AI and IoT-based smart healthcare solutions in urban area[J]. Trends in Health Informatics, 2025, 2(1): 18-26.
- [4] IGE A B, ADEPOJU P A, AKINADE A O, et al. Machine learning in industrial applications: An in-depth review and future directions[J]. International Journal of Multidisciplinary Research and Growth Evaluation, 2025, 6(1): 36-44.
- [5] TIWARI P. The machine learning framework for traffic management in smart cities[J]. Management of Environmental Quality, 2024, 35(2): 445-462.
- [6] 王鑫, 黄伟口, 孙凌云. 跨机构联邦学习的激励机制综述[J]. 计算机科学, 2024, 51(3): 20-29.
- [7] WANG X, HUANG W K, SUN L Y. Survey of incentive mechanism for cross-silo federated learning[J]. Computer Science, 2024, 51(3): 20-29. (in Chinese)
- [8] 陈颖瑜, 李滢东, 张洪磊, 等. 面向可信联邦学习公平性

- 的研究综述[J]. 电子学报, 2023, 51(10): 2985-3010.
- CHEN H Y, LI Y D, ZHANG H L, et al. Fairness in trustworthy federated learning: A survey[J]. *Acta Electronica Sinica*, 2023, 51(10): 2985-3010. (in Chinese)
- [8] RUAN M Z, LI Y H, ZHANG W Z, et al. Optimal power control for over-the-air federated learning with gradient compression[C]//2024 IEEE 30th International Conference on Parallel and Distributed Systems. Piscataway: IEEE, 2024: 326-333.
- [9] MCMAHAN H B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data[C]//Proceedings of the 20th International Conference on Artificial Intelligence and Statistics. Cambridge: PMLR, 2017(54): 1273-1282.
- [10] 马文玉, 陈谦, 胡宇翔, 等. 面向联邦算力物联网的隐私预算自适应优化方案[J]. 物联网学报, 2024, 8(4): 54-69.
- MA W Y, CHEN Q, HU Y X, et al. A privacy budget adaptive optimization scheme for federated computing power Internet of Things[J]. *Chinese Journal on Internet of Things*, 2024, 8(4): 54-69. (in Chinese)
- [11] SUN Y C, SHAO J W, MAO Y Y, et al. Semi-decentralized federated edge learning for fast convergence on non-IID data[C]//2022 IEEE Wireless Communications and Networking Conference. Piscataway: IEEE, 2022: 1898-1903.
- [12] TIAN M M, CHEN Y X, LIU Y, et al. A contract theory based incentive mechanism for federated learning[EB/OL]. (2021-08-12)[2025-09-20]. <https://arXiv.org/abs/2108.05568>.
- [13] LIM W Y B, HUANG J Q, XIONG Z H, et al. Towards federated learning in UAV-enabled Internet of vehicles: A multi-dimensional contract-matching approach[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 22(8): 5140-5154.
- [14] JIAO Y T, WANG P, NIYATO D, et al. Toward an automated auction framework for wireless federated learning services market[J]. *IEEE Transactions on Mobile Computing*, 2021, 20(10): 3034-3048.
- [15] 杜辉, 李卓, 陈昕. 基于在线双边拍卖的分层联邦学习激励机制[J]. 计算机科学, 2022, 49(3): 23-30.
- DU H, LI Z, CHEN X. Incentive mechanism for hierarchical federated learning based on online double auction[J]. *Computer Science*, 2022, 49(3): 23-30. (in Chinese)
- [16] ZHANG J W, WU Y Z, PAN R. Auction-based ex-post-payment incentive mechanism design for horizontal federated learning with reputation and contribution measurement[EB/OL]. (2022-03-15)[2025-09-20]. <https://arXiv.org/abs/2201.02410>.
- [17] LI C L, SONG M Y, LUO Y L. Federated learning based on Stackelberg game in unmanned-aerial-vehicle-enabled mobile edge computing[J]. *Expert Systems with Applications*, 2024, 235: 121023.
- [18] DANG Q L, YANG S, LIU Q Q, et al. Adaptive and communication-efficient zeroth-order optimization for distributed Internet of Things[J]. *IEEE Internet of Things Journal*, 2024, 11(22): 37200-37213.
- [19] WEIBULL J W. *Evolutionary Game Theory*[M]. Cambridge: MIT Press, 1997.
- [20] ZUO C H, XU P H, SONG Y C, et al. RAIM: Three-stage Stackelberg game for hierarchical federated learning with reputation-aware incentive mechanism[J]. *Scientific Reports*, 2025, 15: 34344.
- [21] 康海燕, 冀珊珊. 面向无线边缘网络的分层 Stackelberg 博弈群体激励方法[J]. 电子学报, 2024, 52(7): 2382-2392.
- KANG H Y, JI S S. Hierarchical Stackelberg game swarm learning incentive method for wireless edge network[J]. *Acta Electronica Sinica*, 2024, 52(7): 2382-2392. (in Chinese)
- [22] 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.
- [23] 田有亮, 吴柿红, 李沓, 等. 基于激励机制的联邦学习优化算法[J]. 通信学报, 2023, 44(5): 169-180.
- TIAN Y L, WU S H, LI T, et al. Federated learning optimization algorithm based on incentive mechanism[J]. *Journal on Communications*, 2023, 44(5): 169-180. (in Chinese)
- [24] LIM W Y B, NG J S, XIONG Z H, et al. Decentralized edge intelligence: A dynamic resource allocation framework for hierarchical federated learning[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2022, 33(3): 536-550.
- [25] GOLDSMITH A. *Wireless Communications*[M]. Cambridge: Cambridge University Press, 2005.
- [26] ZHANG Z, JIANG J, XU H Y, et al. Distributed dynamic task allocation for unmanned aerial vehicle swarm systems: A networked evolutionary game-theoretic approach[J]. *Chinese Journal of Aeronautics*, 2024, 37(6): 182-204.
- [27] HOFBAUER J, SIGMUND K. Evolutionary game dynamics[J]. *Bulletin of the American Mathematical Society*, 2003, 40(4): 479-519.
- [28] IMPAGLIAZZO R, LOVETT S, PATURI R, et al. 0-1 in-

teger linear programming with a linear number of constraints[EB/OL]. (2014-02-18)[2025-10-14]. <https://arXiv.org/abs/1401.5512>.

- [29] SAM M L, SAPTARI A, SALLEH M R B, et al. Comparison between linear programming and integer linear programming: A review[J]. *International Journal of Mathematics in Operational Research*, 2018, 13(1): 91.
- [30] DEB K, PAL K. Efficiently solving: A large-scale integer linear program using a customized genetic algorithm[C]// *Genetic and Evolutionary Computation-GECCO 2004*. Berlin: Springer, 2004: 1054-1065.
- [31] 耿方兴, 李卓, 陈昕. 基于多领导者 Stackelberg 博弈的分层联邦学习激励机制设计[J]. *计算机应用*, 2023, 43(11): 3551-3558.

GENG F X, LI Z, CHEN X. Incentive mechanism design for hierarchical federated learning based on multi-leader Stackelberg game[J]. *Journal of Computer Applications*, 2023, 43(11): 3551-3558. (in Chinese)

- [32] KRIZHEVSKY A. Learning multiple layers of features from tiny images[EB/OL]. (2009-04-08)[2025-09-10]. <http://www.cs.utoronto.ca/~kriz/learning-features-2009-TR.pdf>.
- [33] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [34] CHRABASZCZ P, LOSCHILOV I, HUTTER F. A downsampled variant of ImageNet as an alternative to the CIFAR datasets[EB/OL]. (2017-08-23)[2025-09-20]. <https://arXiv.org/abs/1707.08819>.

### 作者简介



刘雨欣 女, 2003年6月出生于四川省德阳市. 现为南京邮电大学通信与信息工程学院本科生. 主要研究方向为联邦学习激励机制.  
E-mail: b22010703@njupt.edu.cn



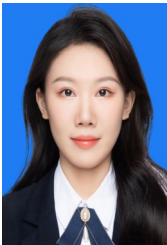
夏文超 男, 1991年4月出生于江苏省盐城市. 现为南京邮电大学通信与信息工程学院副教授. 主要研究方向为通感算协同技术、群智协同技术等.  
E-mail: xiawenchao@njupt.edu.cn



王一航 男, 2004年3月出生于江苏省苏州市. 现为南京邮电大学本科生. 主要研究方向为通信工程.  
E-mail: p22000407@njupt.edu.cn



赵海涛 男, 1983年10月出生于山东省潍坊市. 现为南京邮电大学物联网学院教授. 主要研究方向为智能网络技术、多信道建模技术、物联网、边缘计算等. 中国电子学会会员编号: E190022504S.  
E-mail: zhaopt@njupt.edu.cn



王思洋 女, 2000年2月出生于辽宁省沈阳市. 现为南京邮电大学通信与信息工程学院博士研究生. 主要研究方向为联邦学习.  
E-mail: 2024010101@njupt.edu.cn



卜宪德 男, 1978年10月出生于江苏省徐州市. 现为中国电力科学研究院有限公司正高级工程师. 主要研究方向为电力无线接入网络、异构无线网络通信技术等.  
E-mail: buxiande2@epri.sgcc.com.cn