控制与决策
Control and Decision

# 面向去中心化的零知识联邦半监督学习

陈思光, 潘沭伽

引用本文:
陈思光, 潘沭伽. 面向去中心化的零知识联邦半监督学习[J]. 控制与决策, 2026, 41(5): 1449-1456.

在线阅读 View online: https://doi.org/10.13195/j.kzyjc.2025.0599

---

# 面向去中心化的零知识联邦半监督学习

陈思光[1,2†]，潘沭伽[3]

(1. 河海大学 计算机与软件学院，南京 211100；
2. 南京大学 计算机软件新技术全国重点实验室，南京 210023；
3. 南京邮电大学 物联网学院，南京 210003)

**摘　要:** 为了在去中心化联邦学习场景中实现隐私保护与半监督训练的高效协同, 提出一种面向去中心化的零知识联邦半监督学习算法. 具体地, 首先设计一种反映本地数据特征的零知识特征码, 通过融合 Pedersen 承诺与 Schnorr 证明, 该特征码在实现客户端特征共享的同时, 可保障本地数据不可恢复性与交换过程的合法性验证. 其次, 设计一种高效的去中心化零知识标签传播方法, 利用特征码之间的相似度引导伪标签生成, 在保护隐私的前提下实现高效的标签信息传播, 并通过复杂度分析验证其计算开销显著低于同态加密方案. 最后, 通过在多个数据集上的实验表明, 所提出的算法在不同数据分布与无标签数据配置下均优于现有基准方法, 在准确率和鲁棒性方面具有显著提升; 同时, 通过可变聚类核心数量和网络拓扑结构的实验分析, 进一步验证聚类核心数量对性能的影响, 以及算法在不同去中心化设置中的稳健性和实用性.

**关键词:** 联邦学习；去中心化系统；半监督学习；零知识证明；隐私计算；分布式协作

**中图分类号:** TP181　　**文献标志码:** A

**DOI:** 10.13195/j.kzyjc.2025.0599

**引用格式:** 陈思光, 潘沭伽. 面向去中心化的零知识联邦半监督学习 [J]. 控制与决策, 2026, 41(5): 1449-1456.

# Federated semi-supervised learning with zero-knowledge for decentralized network

*CHEN Si-guang*[1,2†]，*PAN Shu-jia*[3]

(1. College of Computer Science and Software Engineering，Hohai University，Nanjing 211100，China；2. State Key Laboratory for Novel Software Technology，Nanjing University，Nanjing 210023，China；3. School of Internet of Things，Nanjing University of Posts and Telecommunications，Nanjing 210003，China)

**Abstract:** To achieve efficient collaboration between privacy preservation and semi-supervised training in the decentralized federated learning scenario, this paper proposes a decentralized federated semi-supervised learning with zero-knowledge (DFedSem-ZK) algorithm. The algorithm first designs a zero-knowledge feature code that captures local data feature. By integrating Pedersen commitments with Schnorr proofs, the feature code enables secure feature sharing among clients while ensuring the irrecoverability of local data and the verifiability of the exchange process. Furthermore, we construct an efficient decentralized zero-knowledge label propagation method, which leverages the similarity between feature codes to guide pseudo-label generation. This allows for effective dissemination of label information under strict privacy constraints. Computational complexity analysis demonstrates that the computational overhead of the proposed method is significantly lower than that of homomorphic encryption-based schemes. Extensive experiments conducted on multiple datasets show that the algorithm consistently outperforms existing baselines across varying data distributions and unlabeled data configurations, with notable gains in both accuracy and robustness. Additionally, experimental evaluations on the variable number of clustering cores and network topologies further demonstrate the influence of clustering core selection on model performance, as well as the stability and practicality of the proposed algorithm under different decentralized settings.

**Keywords:** federated learning；decentralized systems；semi-supervised learning；zero-knowledge proofs；privacy computing；distributed collaboration

**收稿日期:** 2025-06-06；**录用日期:** 2025-09-24.

**基金项目:** 中央高校基本科研业务费专项资金项目 (B250201047)；江苏省“333 高层次人才培养工程”项目；南京大学计算机软件新技术全国重点实验室开放课题 (KFKT2025B03).

†通信作者. E-mail: sgchen@hhu.edu.cn.

## 0 引 言

机器学习训练需要大量的训练数据, 但随着隐私保护法律法规的发布和用户隐私意识的增强, 高质量数据获取难度和成本持续上升, 如何在保护隐私的同时实现高效的训练成为亟需解决的关键问题[1].

联邦学习 (FL)[2] 通过本地训练并上传参数实现了协作建模, 在一定程度上缓解了隐私问题. 然而, 在现实场景中, 客户端往往只有少量存在标签数据, 而无标签数据大量存在[3], 依赖高质量标注的传统全监督 FL 难以扩展到大规模场景. 联邦半监督学习 (FSSL) 因此诞生, 结合 FL 与 SSL, 使客户端能够在不共享原始数据的前提下充分利用无标签数据提升性能[4].

一种直接的 FSSL 方法是将半监督学习方法 (如 FixMatch[5] 和 FlexMatch[6]) 引入 FL, 但性能提升有限. 针对这种简单组合的不足, 研究者们又进行了一系列改进. 文献 [7] 为伪标签添加了可全局平衡的类置信度阈值, 以实现对每个类别的公平训练; 文献 [8-9] 在有标签数据之间建立类间损失, 通过类间关系指导无标签客户端的训练; 类似地, 文献 [10] 利用样本的类别比例信息辅助监督客户端的训练. 然而, 当数据标签属于非独立同分布 (Non-IID) 时, 这些基于无标签数据进行辅助训练的方法性能会受到较大的不利影响[11].

为解决 Non-IID 带来的负面影响, RSCFed 通过随机子采样和加权聚合提升无标签数据利用率[12]. 文献 [13] 针对客户端数据完全无标签的情况, 提出了交替训练策略, 即服务器使用少量标签数据微调全局模型并生成伪标签供客户端训练. FL 通过锐度感知一致性正则化改进无标签训练[14]. 文献 [15] 结合高低置信度样本缓解类别不平衡并提升泛化能力. 这些工作虽提升了性能, 但主要集中于训练优化, 在隐私和安全性方面仍缺乏有效设计.

现有的 FSSL 方法大多依赖于中心服务器进行聚合, 但在更贴近现实环境的去中心化场景下的可部署性和有效性仍面临挑战[16]. 部分研究提出点对点协作, 如基于邻居选择的去中心化学习[17]. 文献 [18] 提出了一种新的去中心化半监督学习方法, 针对有限的去中心化医学图像和部分标注数据进行了研究; 文献 [19] 设计了一个基于共识的扩散模型来生成合成数据, 以充分利用去中心化网络中的邻域信息. 然而, 伪标签生成机制仍可能泄露类别分布或本地偏好, 缺乏验证和防御机制. 部分研究者探索引入差分隐私[20] 或客户端选择策略[21] 抵御攻击. 相关研究仍停留在“可用性”层面, 而在“安全性和可验证性”上存在研究空白. 如何在去中心化场景下兼顾半监督训练的效率和零知识安全保障, 是本文要解决的核心问题.

为了实现高效且安全的联邦半监督学习, 并进一步提升其在去中心化异构环境下的实用性和鲁棒性, 本文提出一种面向去中心化场景的零知识联邦半监督学习算法 (DFedSem-ZK), 重点解决现有方法在隐私泄露、投毒攻击防御不足、缺乏可验证性等方面的问题. 具体贡献如下:

1) 融合 Pedersen 承诺[22] 与 Schnorr 证明[23], 设计一种不可恢复与可验证的零知识特征码, 用于表示本地数据特征, 在实现特征共享的同时保护本地数据偏好并保证可验证性.

2) 设计一种高效的去中心化零知识标签传播方法, 基于特征码相似度加权传播生成高质量伪标签, 并结合正则化约束的加权损失开展模型训练, 实现全局性能提升, 开销显著低于同态加密.

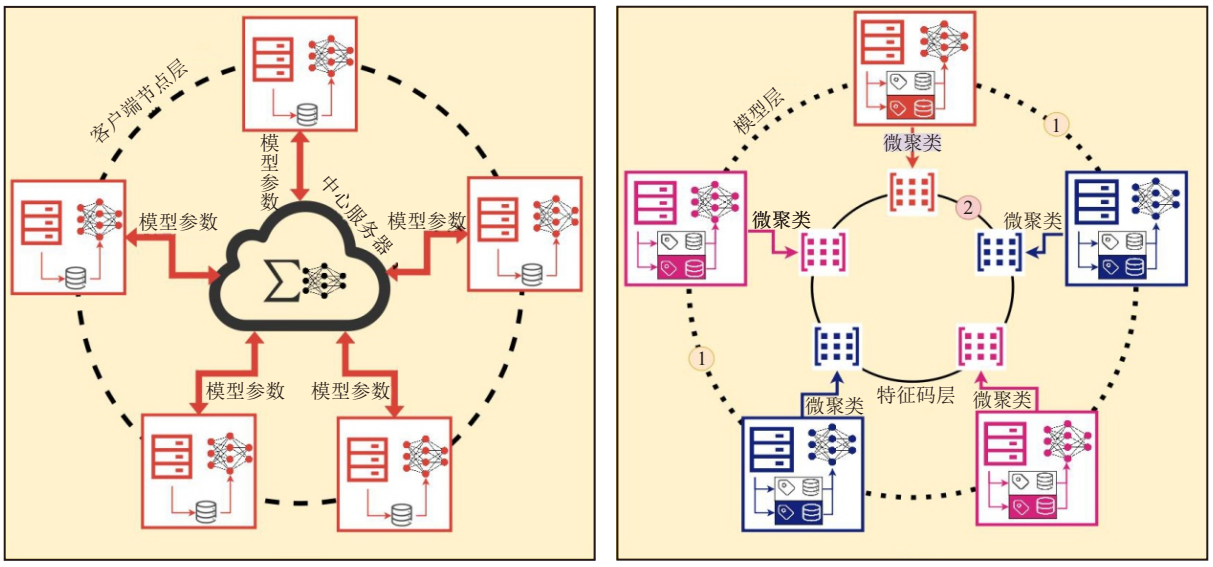
3) 通过实验验证算法在异构分布和标签稀缺场景下的有效性, 分析其在不同拓扑结构和聚类规模下的适应性. 结果表明, 相比多种现有 FSSL 方法, 本文算法在准确率和鲁棒性上均有提升.

## 1 网络模型

传统联邦学习依赖中心服务器协调全局模型聚合, 如图 1(a) 所示. 客户端在本地数据集上训练后, 将模型参数或梯度上传至中央服务器, 由其完成聚合并下发更新模型. 本文考虑的去中心化联邦学习场景中, 客户端构成全连接网络, 不再依赖中心节点. 基于该场景, 本文设计一个由模型层与特征码层组成的隐私保护型联邦半监督学习框架, 如图 1(b) 所示.

1) 模型层: 模型层由可以进行点对点通信的客户端节点构成, 核心功能包括两个方面. 首先是特征提取与聚类: 客户端利用预训练网络提取本地数据特征, 并进行聚类, 生成特征聚类集合作为特征码, 用于表示标签在高维空间中的分布. 这一过程既减少了数据量, 又提升了特征表达质量, 为全局聚类提供高质量输入. 其次是半监督训练与参数更新: 客户端根据特征码相似度生成伪标签, 并基于融合损失进行半监督训练. 训练过程中, 节点通过参数交换完成去中心化聚合, 有效利用有标签数据推动伪标签生成和模型更新.

2) 特征码层: 负责实现隐私保护下的特征码交




(a)　传统联邦学习　　(b)　去中心化联邦半监督学习

**图1　网络模型**

换. 为防止特征码被逆向恢复, 客户端先对其加密生成零知识特征码. 在首次通信中, 客户端通过零知识证明识别并排除恶意节点. 验证通过后, 客户端根据相似度迭代更新本地特征码, 使其更好地表征本地特征, 同时避免与邻居差异过大. 最终, 客户端在特征码收敛后采用带原文约束的标签传播逐步扩散标签, 实现全局一致性. 该设计既能够保障隐私安全, 又能够确保标签传播的可靠性和效率.

## 2　去中心化的零知识联邦半监督学习

本部分提出一种去中心化零知识联邦半监督学习算法, 结合零知识特征码与去中心化标签传播, 实现数据隐私保护与半监督联邦学习的融合. 如图 2 所示, 整体框架包括零知识特征码生成、去中心化零知识标注和基于标注标签的联邦半监督学习 3 部分.

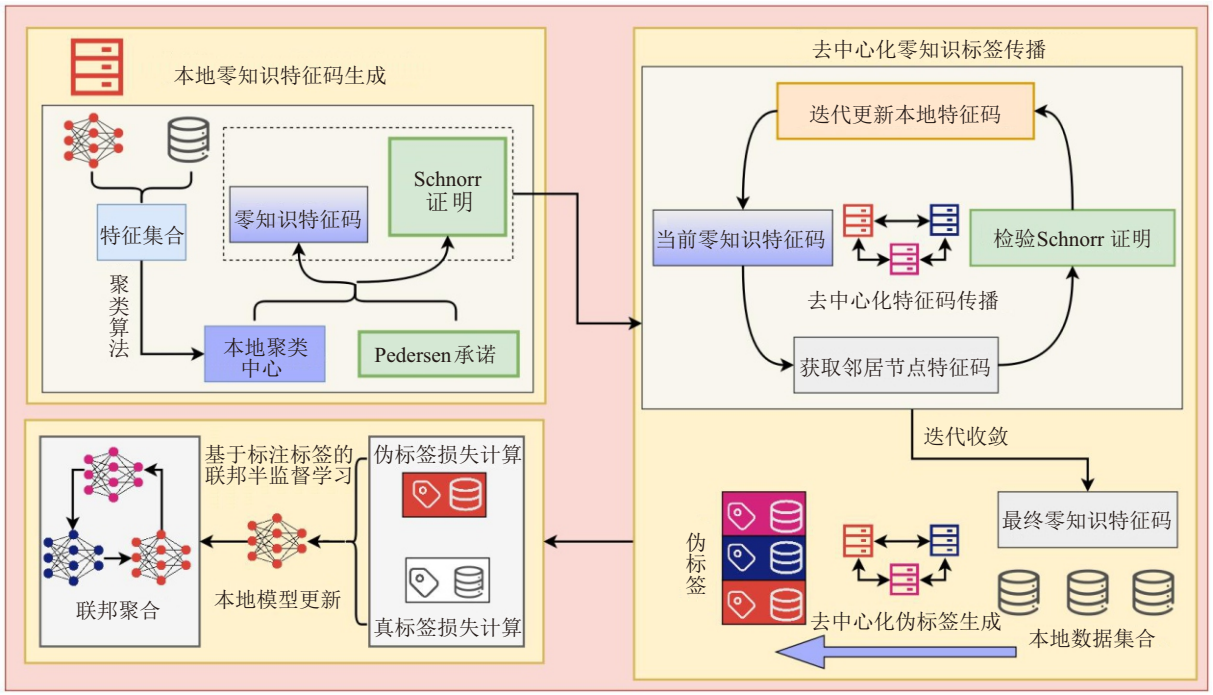



**图2　去中心化的零知识联邦半监督学习框架**

### 2.1　零知识特征码

本节设计基于本地数据特征向量的零知识特征码, 通过本地特征聚类集合和 Pedersen 承诺, 对聚类核心集合加密后生成隐私保护的零知识特征码, 确保聚类核心集合在交换和共享过程中的不可恢复性.

对于客户端 $k$, 其本地数据集$D_k$包含$n_k$个样本$[x_1, x_2, \ldots, x_{n_k}]$. 客户端基于预训练模型$\theta$对每个样本进行特征提取, 生成特征向量$f_i$, 如下所示:

$$f_i = \phi(x_i; \theta), \tag{1}$$

其中特征提取函数记为$\phi(\,;\theta)$. 客户端 $k$ 的特征向量集合$F_k = \{f_i \,|\, f_i = \phi(x_i;\theta), x_i \in D_k\}$. 接下来采用基于小样本采样的 MiniBatch $K$-Means 算法, 将特征集合$F_k$聚类为$N_k$个簇, 生成聚类核心集合$C_k$, 降低特征维度并提取关键信息

$$C_k = \text{MiniBatchKMeans}(F_k, N_k), \tag{2}$$

其中$N_k$为客户端 $k$ 预设的聚类核心数目.

为了完整地保证聚类核心的隐私性并在后续的全局聚类过程中达成零知识, 客户端将$C_k$通过 Pedersen 承诺机制进行域迁移. 首先为每个聚类核心簇$C_{k,i}$生成随机数$r_i$, 然后根据式 (3) 计算加密后的聚类核心集合, 即零知识特征码.

$$\tilde{C}_k = \{g^{r_i} h^{C_{k,i}} \,|\, i = 1, 2, \ldots, N_k\}. \tag{3}$$

其中: $g$和 $h$ 为预定义的群元素, 长度可满足离散对数难以求解的难题. 客户端后续共享的内容变为零知识特征码$\tilde{C}_k$, 而非聚类核心集合$C_k$.

### 2.2 去中心化零知识标签传播

在完成初始本地零知识特征码$\tilde{C}_k^0$的计算后, 每个客户端会与其邻居客户端$j \in J_k$($J_k$是客户端 $k$ 的邻居集合) 进行多轮通信, 吸收其他客户端的特征码信息并逐步更新本地特征码

$$\tilde{C}_k^{t+1} = \tilde{C}_k^t \prod_{j \in J_k} \left(\frac{\tilde{C}_j^t}{\tilde{C}_k^t}\right)^{\alpha_{k,j}}, \tag{4}$$

其中$\alpha_{k,j}$为加权系数.

在第 1 轮通信时, 客户端将提供对应的 Schnorr 证明, 确保$C_k$确实是在定义域内, 并且$\tilde{C}_k^0$是根据 Pedersen 承诺正确构造的. 具体证明过程如下: 客户端 $k$ 选择随机数集合$s = \{s_i \,|\, i = 1, 2, \ldots, N_k\}$, 通过公式计算承诺值$A = g^s$, 接受来自相邻节点的挑战值集合$e$, 根据$e$依次计算响应值

$$p = s + eC_k, \tag{5}$$

通过验证方程$g^p = A\tilde{C}_k^{0e}$完成证明. 这一机制不仅保证了数据的合法性, 还确保在验证过程中不会泄露任何额外信息.

由于 Pedersen 具有加法同态性质, $\tilde{C}_k$仍然能够保留与$C_k$一致的分布特征, 结合 Pedersen 承诺的过程进行推导, 式 (4) 的迭代过程可推导至

$$\begin{aligned}
&\tilde{C}_k^{t+1} = \tilde{C}_k^t \prod_{j \in J_k} \left(\frac{\tilde{C}_j^t}{\tilde{C}_k^t}\right)^{\alpha_{k,j}} = \\
&\tilde{C}_k^t g^{\sum\limits_{j \in J_k} (r_j - r_k)\alpha_{k,j}} h^{\sum\limits_{j \in J_k} (C_j^t - C_k^t)\alpha_{k,j}} = \\
&g^{r_k} h^{C_k^t} g^{\sum\limits_{j \in J_k} (r_j - r_k)\alpha_{k,j}} h^{\sum\limits_{j \in J_k} (C_j^t - C_k^t)\alpha_{k,j}} = \\
&g^{r_k + \sum\limits_{j \in J_k} (r_j - r_k)\alpha_{k,j}} h^{C_k^t + \sum\limits_{j \in J_k} (C_j^t - C_k^t)\alpha_{k,j}},
\end{aligned} \tag{6}$$

$$C_k^{t+1} = C_k^t + \sum_{j \in J_k} (C_j^t - C_k^t)\alpha_{k,j}. \tag{7}$$

随着迭代轮次 $t$ 的增加, $\tilde{C}_k^t$基于与邻居节点特征码的距离加权和进行迭代, $\tilde{C}_k^t$逐渐迭代至$\tilde{C}_k^T$.

当特征码信息扩散的过程完成时, 基于特征相似性度量的原文约束无标签标注将开展. 首先, 计算客户端$k$与其邻居客户端$j$对应的每对簇的特征码相似度, 如下所示:

$$\begin{aligned}
&W_{\tilde{C}_{k,i}^T, \tilde{C}_{j,i}^T} = \\
&\exp\Big(-\frac{\|\tilde{C}_{k,i}^T / \tilde{C}_{j,i}^T\|_2^2}{2\sigma^2}\Big) \exp(-\lambda \|C_{k,i}^0 - \varphi(\tilde{C}_{k,i}^T)\|_2^2) \cdot \\
&\exp(-\lambda \|C_{j,i}^0 - \varphi(\tilde{C}_{j,i}^T)\|_2^2).
\end{aligned} \tag{8}$$

其中: $\tilde{C}_{k,i}^T$和$\tilde{C}_{j,i}^T$为信息传播后, 客户端$k$和$j$特征码的第$i$个簇; $\sigma$为相似性度量中的超参数; $\exp\Big(-\frac{\|\tilde{C}_{k,i}^T / \tilde{C}_{j,i}^T\|_2^2}{2\sigma^2}\Big)$通过计算客户端$k$和$j$特征码的第$i$个簇的欧几里得距离, 对客户端之间每个簇$i$的相似度进行衡量. 客户端$k$和$j$通过$\varphi(\cdot)$对特征码的内容解密, 计算簇$i$相较原始聚类核心集合的变化量并发送给对方, 对原始距离进行约束.

接下来$k$的簇$i$会被赋予一个初始标签$\tilde{y}_{k,i}^0$, 即

$$\tilde{y}_{k,i}^0 = \arg\max_{\gamma} \sum_{\phi(x_z;\theta) \in \tilde{C}_{k,i}^T} 1(y_z = \gamma). \tag{9}$$

其中: $x_z$为属于$\tilde{C}_{k,i}^T$的特征所对应的数据样本, 当$x_z$有标签且标签$y_z$为$\gamma$时, 计数加 1, 找到拥有最多数据样本类别$\gamma$, 作为簇$i$的初始标签.

接下来, 通过下式利用相似度矩阵$W$对每个簇进行标签传播:

$$\tilde{y}_{k,i}^{t+1} = \arg\max \Big(\sum_{j \in J_k} W_{C_{k,i}^T, C_{j,i}^T} \cdot \tilde{y}_{j,i}^t + \tilde{y}_{k,i}^t\Big). \tag{10}$$

其中: $\tilde{y}_{k,i}^t$将与其邻居节点的标签一起, 作为软标签加权求和. 每一轮迭代都取出软标签中概率最大的标签, 将其作为下一轮的初始值. 随着$t$的增加, 标签会在全网客户端之间逐渐扩散, $\tilde{y}_{k,i}^t$将逐渐收敛, 客户端$k$的簇$i$都会得到自己的标签, 即簇$i$内的无标签数据会被标记为其所属的$\tilde{y}_{k,i}^T$.

### 2.3 基于标注标签的联邦半监督学习

本文方法本地训练的目标分为两部分: 有标签

数据使用真实的标注数据$(x_i, y_i)$进行监督学习; 无标签数据的标签为其所属簇$i$最终的标签$\tilde{y}_{k,i}^T$, 为简化表示, 记为$\tilde{y}_i$. 本地模型的总体目标函数为最小化有标签数据的监督损失$\mathcal{L}_{\text{sup}}$与无标签数据的伪标签损失$\mathcal{L}_{\text{unsup}}$之和.

对于有标签的数据, 使用标准的监督学习目标函数. 假设客户端$k$拥有$n_l$个标注数据$\{(x_i, y_i)\}_{i=1}^{n_l}$, 其监督损失定义为

$$\mathcal{L}_{\text{sup}} = \frac{1}{n_l}\sum_{i=1}^{n_l} l(y_i, f(x_i;\omega_k)). \tag{11}$$

其中: $f(x_i;\omega_k)$为本地模型在输入$x_i$下的预测结果; $l$为损失函数, 用于衡量预测结果与真实标签之间的距离.

对于无标签的数据, 使用伪标签$\tilde{y}_i$作为监督信号. 假设客户端$k$拥有$n_u$个无标签数据$\{(x_i)\}_{i=1}^{n_u}$, 其伪标签损失定义为

$$\mathcal{L}_{\text{unsup}} = \frac{1}{n_u}\sum_{i=1}^{n_u}[l(\tilde{y}_i, f(x_i;\omega_k)) + \|f(x_i;\omega_k) - f(x_i';\omega_k)\|_2^2]. \tag{12}$$

其中: $\|f(x_i;\omega_k) - f(x_i';\omega_k)\|_2^2$为一致性正则化项, $x_i'$为$x_i$的不同视图.

每轮训练得到损失后, 客户端$k$以$\eta$为学习率更新本地模型参数, 再通过联邦聚合平均的方式更新本地模型.

### 2.4　复杂度分析

明文聚类核心集合$C_k$的场景下, 特征提取和聚类的计算复杂度为$O(n_k N_k d)$. 其中: $n_k$为数据样本数量, $N_k$为聚类核心数, $d$为特征维度. 特征码传播阶段的计算量主要涉及加法运算, 复杂度约为$O(T_1 K \tilde{J})$. 其中: $T_1$为传播轮数, $K$为客户端总数, $\tilde{J}$为每个客户端的平均邻居数. 在标签传播过程中, 复杂度分别为$O(T_2 K \tilde{J})$和$O({N_k}^2 K)$, $K$为客户端总数, $T_2$为传播轮数.

本文方法在初始聚类阶段, 每个客户端需要对$N_k$个聚类核心进行承诺, 且计算复杂度为$O(n_k N_k d \log q)$, $q$为群的阶数. 在特征码传播过程中, 由于 Pedersen 承诺的加法同态性, 零知识的特征码更新通过乘除完成, 但复杂度仍然为$O(T_1 K \tilde{J})$. 在计算$\varphi(\tilde{C}_{k,i}^T)$时, 需要求离散对数, 采用 Pollard-Rho 算法解密时, 复杂度约为$O({N_k}^2 K\sqrt{q})$.

如果将 Pedersen 承诺改为无损的同态加密方案 Paillier[24], 则初始加密阶段, 每个客户端执行加密的复杂度上升至$O(n_k N_k d \log q^2)$, 而零知识特征码传播时, 每次更新都涉及指数运算, 复杂度增至$O(T_1 K \tilde{J} \log q^2)$. 此外, 在计算相似度矩阵$W$时, 执行涉及到模幂运算, 复杂度为$O(N_k^2 K \log q^3)$.

整体而言, 本文方案与明文的计算开销在同一数量级内, Paillier 方案则多出 2 或 3 个数量级的计算开销. 本文方法采用 Pedersen 承诺与正则项结合的方法, 在复杂度方面显著优于同态加密方案.

## 3　实　验

本节通过一系列实验验证所提出的 DFedSem-ZK 算法在不同场景下的有效性和性能优势, 分析聚类核心数量对收敛的影响, 研究不同网络拓扑下的模型准确率.

### 3.1　实验设置

数据集与数据划分: 实验以 CIFAR-10, CIFAR-100 和 Mini-ImageNet 作为数据集. CIFAR-10 设置为 75% 与 50% 无标签比例, CIFAR-100 和 Mini-ImageNet 分别为 50% 和 25%. 在此基础上设计 IID 和狄利克雷分布的 Non-IID 两种场景. 此外, 对无标签数据的分布设计集中在部分客户端 (50%K) 和分散在所有客户端 (100%K) 这两种场景.

基准方法: 对比包括 Fed-UDA[25]、FedSem+[26]、Fed-PL[27]、FedMatch[28]、Fed-FixMatch[5]和 FedSiam[29]方法. 实验在 IID 场景下使用 FedAvg 作为基座方法, 在 Non-IID 场景下使用 FedProx[30]. Fed-PL 和 FedSem+都依赖于伪标签进行半监督训练, Fed-FixMatch、FedSiam 和 FedMatch 采用一致性正则化方法.

具体配置: CIFAR-10 的实验中使用 ResNet-9 作为本地分类模型, CIFAR-100 和 Mini-ImageNet 则使用 ResNet-18. MiniBatch $K$-Means 的聚类核心数设为类别数乘$e$. 准确率实验取两次随机结果平均, 置信区间分别为 CIFAR-10: IID 0.4%、Non-IID 2.5%; CIFAR-100 和 Mini-ImageNet: 1.5% 和 4.8%. Non-IID 实验中 Dirichlet 参数$\alpha$为 0.3. 实验代码在 RTX4090 显卡上使用 PyTorch 搭建.

### 3.2　准确率实验

如表 1 和表 2 所示, DFedSem-ZK 在 CIFAR-10 和 CIFAR-100 上均取得了显著提升. 在 CIFAR-10 的 IID 场景下, 当标注数据仅为 15 000 时, DFedSem-ZK 的准确率可达 71.6%, 显著高于伪标签方法 Fed-PL (65.2%) 和基于正则化的方法 FedSiam (62.9%), 体现了零知识特征码在特征提取和聚类中的优势. 借助 MiniBatch $K$-Means 聚类和 Pedersen 承诺保护, DFedSem-ZK 无需暴露原始数据即可高效标注. 尤其在 50%K 场景下, 依然表现稳健, 明显优于基于伪标签和正则化的方法, 表明其对标注数据分布不均

**表1 CIFAR-10 准确率实验**

| 方法 | CIFAR-10 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 75%无标签数据 | | | | 50%无标签数据 | | | |
| | 50%K | | 100%K | | 50%K | | 100%K | |
| | IID | Dirichlet | IID | Dirichlet | IID | Dirichlet | IID | Dirichlet |
| Fed-UDA | 47.5 | 46.31 | 43.5 | 44.6 | 65.3 | 64.1 | 63.0 | 61.9 |
| Fed-FixMatch | 47.2 | 45.5 | 46.4 | 45.1 | 64.4 | 63.0 | 63.0 | 60.5 |
| Fed-PL | 65.2 | 60.3 | 61.7 | 50.9 | 84.1 | 78.3 | **85.1** | 76.6 |
| FedSem+ | 59.5 | 59.7 | 60.7 | 61.0 | 80.9 | 81.3 | 80.1 | 79.8 |
| FedSiam | 62.9 | 62.3 | 67.3 | 67.8 | 78.4 | 78.8 | 81.0 | 82.2 |
| FedMatch | 57.1 | 48.2 | 50.9 | 50.7 | 77.8 | 66.3 | 72.2 | 72.0 |
| DFedSem-ZK | **71.6** | **69.5** | **69.5** | **68.1** | **85.5** | **82.6** | 84.2 | **83.0** |

**表2 CIFAR-100 和 Mini-ImageNet 准确率实验**

| 方法 | CIFAR-100 | | | | Mini-ImageNet | | | |
|---|---|---|---|---|---|---|---|---|
| | IID | | | | IID | | | |
| | 50%无标签数据 | | 25%无标签数据 | | 50%无标签数据 | | 25%无标签数据 | |
| | 50%K | 100%K | 50%K | 100%K | 50%K | 100%K | 50%K | 100%K |
| Fed-PL | 45.8 | 44.5 | 57.5 | 55.3 | 25.5 | **25.3** | **34.8** | 33.0 |
| FedSem+ | 46.2 | 45.6 | 55.2 | 53.8 | 24.8 | 23.9 | 32.5 | 32.2 |
| FedSiam | 45.1 | 44.4 | 56.2 | 56.4 | 25.2 | 23.4 | 34.5 | **33.6** |
| FedMatch | 41.3 | 39.9 | 48.3 | 46.0 | 22.3 | 20.8 | 29.6 | 31.8 |
| DFedSem-ZK | **50.6** | **49.6** | **59.3** | **58.6** | **25.6** | 24.3 | 34.0 | 33.2 |

具有强适应能力. 在 CIFAR-10 低比例无标签的 Non-IID 场景下, DFedSem-ZK 的 82.6% 准确率同样领先. 这主要得益于去中心化零知识标注的多轮通信机制, 使客户端逐步吸收邻居聚类中心信息, 缓解了数据异构挑战. CIFAR-100 实验中, DFedSem-ZK 同样保持领先: 在 IID 下准确率达 50.6%, 在 10 000 标注样本下达到 59.3% 和 58.6%, 均优于对比方法. 这源于零知识特征码对高维特征的压缩和去中心化标注的信息传播, 使其在复杂任务中依旧稳健. 相比其他方法, DFedSem-ZK 不仅更好地利用了无标签数据, 还结合 Schnorr 证明保障了数据交换的合法性和安全性.

尽管 Mini-ImageNet 类别众多且特征复杂, DFedSem-ZK 依然保持了竞争力: 在 45 K 标注样本的 IID 下准确率为 25.6%, 在 25 K 样本下提升至 33.2%, 与主流方法持平. 这表明 DFedSem-ZK 在高维任务中仍具鲁棒性和通用性.

总体而言, 得益于零知识特征码和去中心化标注的设计, DFedSem-ZK 在 CIFAR-10 和 CIFAR-100 上均展现显著优势, 尤其在标注稀缺、分布不均及任务复杂的场景下表现突出. 虽然在 Mini-ImageNet 上优势有限, 但其稳定表现进一步验证了方法的实用价值.

### 3.3 聚类核心数量实验

图 3 和图 4 的实验结果表明, 聚类的核心数量设置对模型性能的影响不仅与数据集复杂度及分布特性紧密相关, 数据在客户端的分布情况也对其有明显影响. 如图 3 所示, 针对 CIFAR-10 (10 类别), 当核心数较少 (等于类别数) 时, 模型在 IID (图 3(a)) 和 Non-IID (图 3(b)) 场景下均表现过拟合, 但 IID 场景中由于数据分布的均衡性, 模型对局部噪声的敏感性较低, Non-IID 场景下, 客户端间的异构分布放大了局部统计偏差, 因此 IID 场景下的过拟合程度相

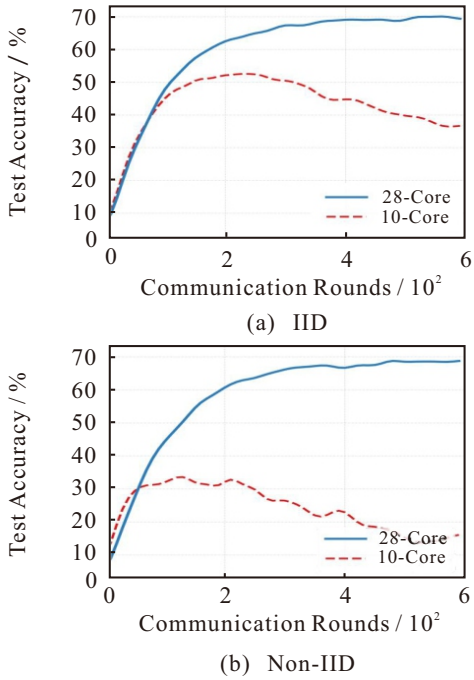
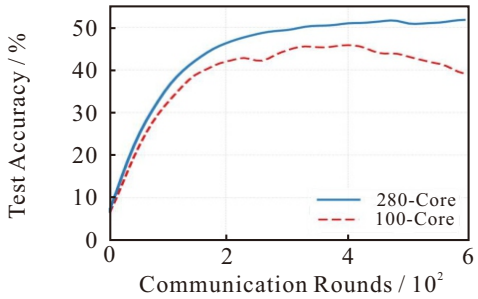
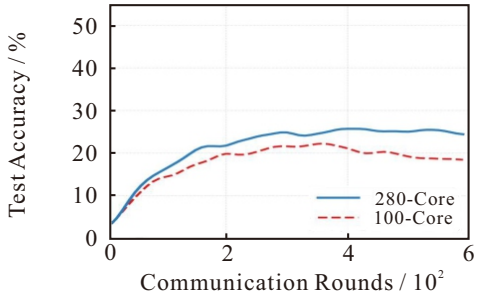



**图3 CIFAR10 中聚类核心数量对收敛的影响**




(a) CIFAR100




(b) Mini-ImageNet

**图4　聚类核心数量对收敛的影响**

对较轻.

对于类别更细粒度的 CIFAR-100 (图 4(a), 100 类别) 和 Mini-ImageNet (图 4(b), 100 类别), 如图 4 所示, 尽管模型仍会表现出过拟合, 但程度相对同等聚类核心数设置下的 CIFAR10 已经得到减轻, 并且随着特征空间复杂度的提升 (CIFAR100 到 Mini-ImageNet), 过拟合得到进一步的缓解, 这是由于特征空间复杂程度的上升弱化了噪声带来的影响. 但由于特征空间复杂度提升, 较少的聚类核心数量导致每个类别的容量不足以区分高维语义差异, 表现为准确率系统性下降, 表明在类别更细粒度的场景下, 欠拟合会成为主要矛盾.

上述结果表明, 聚类复杂度需与数据集特性动态匹配, 低复杂度任务需抑制过拟合, 高复杂度任务则需更多核心或层级表征以避免欠拟合.

### 3.4　网络拓扑对性能的影响

本实验通过调整 Erdos–Rényi 模型中的连接概率, 分析客户端连接密度对性能的影响. 实验在 CIFAR-10、15 000 标注样本、IID 分布下进行.

表 3 结果显示, 连接概率下降会导致准确率下降, 且集中无标签数据 (50%K) 更为敏感: 当连接概率降至 0.3 时, 准确率较全连接下降 18.2%, 而分散分布 (100%K) 仅下降 5.5%. 说明稀疏网络下, 分散策略更具鲁棒性. 综合而言, 连接概率大于 0.8 时接近全连接性能, 而在信道受限场景下需保持至少 0.5, 以避免过度稀疏.

该实验验证了 DFedSem-ZK 在不同网络密度下的适应性. 即便在较低连接概率下, 得益于去中心化零知识标注, 其仍能保持一定性能, 尤其在数据分散场景下更稳健.

**表3　不同客户端连接数量下的准确率实验**

| 连接概率 | 1.0 | 0.8 | 0.5 | 0.3 |
|---|---|---|---|---|
| 50%K | 71.6 | 71.4 (−0.2%) | 66.2 (−7.5%) | 58.6 (−18.2%) |
| 100%K | 69.5 | 69.0 (−0.5%) | 67.9 (−2.3%) | 65.7 (−5.5%) |

## 4　结　论

本文提出的 DFedSem-ZK 面向去中心化场景, 解决了联邦学习中标签稀缺与隐私保护的矛盾. 其核心是不可恢复且可验证的零知识特征码, 结合高效的零知识标签传播, 实现了性能与效率的平衡. 实验结果显示, 该方法在准确率和鲁棒性上均优于现有方案, 并在网络稀疏环境中保持适应性, 同时探索了聚类核心数对性能的影响, 验证了其实用性.

## 参考文献 (References)

[1] Shokri R, Shmatikov V. Privacy-preserving deep learning[C]. Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. New York, 2015: 1310-1321.

[2] McMahan H B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[C]. International Conference on Artificial Intelligence and Statistics. Fort Lauderdale, 2017: 1273-1282.

[3] 艾明曦, 许庆, 张进, 等. 基于局部特征增强的浮选过程改进半监督工况识别方法[J]. 控制与决策, 2025, 40(9): 2891-2900.
(Ai M X, Xu Q, Zhang J, et al. Local feature enhancement-based improved semi-supervised condition recognition method for flotation process[J]. Control and Decision, 2025, 40(9): 2891-2900.)

[4] Shi Y H, Chen S G, Zhang H J. Uncertainty minimization for personalized federated semi-supervised learning[J]. IEEE Transactions on Network Science and Engineering, 2023, 10(2): 1060-1073.

[5] Sohn K, Berthelot D, Carlini N, et al. FixMatch: Simplifying semi-supervised learning with consistency and confidence[C]. Proceedings of the 34th Annual Conference on Neural Information Processing Systems. Virtual Event, 2020: 596-608.

[6] Zhang B, Wang Y, Hou W, et al. FlexMatch: Boosting semi-supervised learning with curriculum pseudo labeling[C]. Proceedings of the 35th Annual Conference on Neural Information Processing Systems. Virtual Event, 2021: 18408-18419.

[7] Wei X X, Huang H. Balanced federated semisupervised learning with fairness-aware pseudo-labeling[J]. IEEE Transactions on Neural Networks and Learning Systems,

2024, 35(7): 9395-9407.

[8] Liu Q D, Yang H Z, Dou Q, et al. Federated semi-supervised medical image classification *via* inter-client relation matching[C]. Medical Image Computing and Computer Assisted Intervention-MICCAI 2021. Cham: Springer International Publishing, 2021: 325-335.

[9] Liu Y Z, Wu H S, Qin J. FedCD: Federated semi-supervised learning with class awareness balance *via* dual teachers[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2024, 38(4): 3837-3845.

[10] Itahara S, Nishio T, Koda Y, et al. Distillation-based semi-supervised federated learning for communication-efficient collaborative training with non-IID private data[J]. IEEE Transactions on Mobile Computing, 2023, 22(1): 191-205.

[11] Zhang C, Wu F Z, Yi J W, et al. Non-IID always bad? Semi-supervised heterogeneous federated learning with local knowledge enhancement[C]. Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. New York, 2023: 3257-3267.

[12] Liang N, Lin Y, Fu H, et al. Federated semi-supervised medical image segmentation via uncertainty-aware self-ensembling model[C]. Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, 2022: 10154-10163.

[13] Enmao D, Jie D, Vahid T. SemiFL: Semi-supervised federated learning for unlabeled clients with alternate training[C]. Proceedings of the 36th Annual Conference on Neural Information Processing Systems. New Orleans, 2022: 17871-17884.

[14] Lee S, Le T L V. FL$^2$: Overcoming few labels in federated semi-supervised learning[C]. Proceedings of the 38th Annual Conference on Neural Information Processing Systems. Vancouver, 2024: 43693-43714.

[15] Chen S G, Shen J H. Exploitation maximization of unlabeled data for federated semi-supervised learning[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2025, 9(2): 2039-2044.

[16] Albaseer A, Abdallah M, Al-Fuqaha A, et al. Semi-supervised federated learning over heterogeneous wireless IoT edge networks: Framework and algorithms[J]. IEEE Internet of Things Journal, 2022, 9(24): 25626-25642.

[17] Jiang Z D, Xu Y, Xu H L, et al. Semi-supervised decentralized machine learning with device-to-device cooperation[J]. IEEE Transactions on Mobile Computing, 2024, 23(10): 9757-9771.

[18] Dong N Q, Kampffmeyer M, Voiculescu I, et al. Federated partially supervised learning with limited decentralized medical images[J]. IEEE Transactions on Medical Imaging, 2023, 42(7): 1944-1954.

[19] Liu X Y, Han P C, Li X, et al. SemiDFL: A semi-supervised paradigm for decentralized federated learning[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2025, 39(18): 18987-18995.

[20] Fan C Y, Hu J J, Huang J W. Private semi-supervised federated learning[C]. Proceedings of the 31st International Joint Conference on Artificial Intelligence. Vienna, 2022: 2009-2015.

[21] Liu Y, Yuan X L, Zhao R H, et al. Poisoning semi-supervised federated learning via unlabeled data: Attacks and defenses[J/OL]. 2020, arXiv: 2012.04432.

[22] Pedersen T P. Non-interactive and information-theoretic secure verifiable secret sharing[C]. Advances in Cryptology — CRYPTO'91. Heidelberg, 1992: 129-140.

[23] Schnorr C P. Efficient identification and signatures for smart cards[C]. Advances in Cryptology — CRYPTO' 89 Proceedings. New York, 1990: 239-252.

[24] Paillier P. Public-key cryptosystems based on composite degree residuosity classes[C]. Advances in Cryptology — EUROCRYPT'99. Berlin, 1999: 223-238.

[25] Xie Q, Dai Z, Hovy E, et al. Unsupervised data augmentation for consistency training[C]. Proceedings of the 34th Annual Conference on Neural Information Processing Systems. Virtual Event, 2020, 6256-6268.

[26] Albaseer A, Ciftler B S, Abdallah M, et al. Exploiting unlabeled data in smart cities using federated edge learning[C]. 2020 International Wireless Communications and Mobile Computing. Limassol, 2020: 1666-1671.

[27] Li M, Li Q L, Wang Y. Class balanced adaptive pseudo labeling for federated semi-supervised learning[C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, 2023: 16292-16301.

[28] Jeong W, Yoon J, Yang E, et al. Federated semi-supervised learning with inter-client consistency & disjoint learning[J/OL]. 2020, arXiv: 2006.12097.

[29] Long Z W, Che L W, Wang Y Q, et al. FedSiam: Towards adaptive federated semi-supervised learning[J/OL]. 2020, arXiv: 2012.03292.

[30] Li T, Sahu A K, Talwalkar A, et al. FedProx: A robust federated learning framework with heterogeneous devices[J/OL]. 2020, arXiv: 1812.06127.

## 作者简介

陈思光 (1984−), 男, 教授, 博士, 主要研究方向为人工智能与安全, E-mail: sgchen@hhu.edu.cn;

潘沭伽 (2000−), 男, 硕士生, 主要研究方向为分布式机器学习, E-mail: panshujia2000@outlook.com.