



# 计算机科学

COMPUTER SCIENCE

## 基于改进NSGA-III的多目标联邦学习进化算法

钟佳淋, 吴亚辉, 邓苏, 周浩浩, 马武彬

引用本文

钟佳淋, 吴亚辉, 邓苏, 周浩浩, 马武彬 [基于改进NSGA-III的多目标联邦学习进化算法](#) [J]. 计算机科学, 2023, 50(4): 333-342.

ZHONG Jialin, WU Yahui, DENG Su, ZHOU Haohao, MA Wubin. [Multi-objective Federated Learning Evolutionary Algorithm Based on Improved NSGA-III](#) [J]. Computer Science, 2023, 50(4): 333-342.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

### [演化循环神经网络研究综述](#)

Survey on Evolutionary Recurrent Neural Networks

计算机科学, 2023, 50(3): 254-265. <https://doi.org/10.11896/jsjcx.220600007>

### [基于李雅普诺夫优化的移动群智感知在线任务分配策略](#)

Online Task Allocation Strategy Based on Lyapunov Optimization in Mobile Crowdsensing

计算机科学, 2023, 50(2): 50-56. <https://doi.org/10.11896/jsjcx.221100179>

### [基于联邦学习的暖通空调系统故障检测与诊断](#)

Fault Detection and Diagnosis of HVAC System Based on Federated Learning

计算机科学, 2022, 49(12): 74-80. <https://doi.org/10.11896/jsjcx.220700280>

### [基于联邦学习的Gamma回归算法](#)

FL-GRM: Gamma Regression Algorithm Based on Federated Learning

计算机科学, 2022, 49(12): 66-73. <https://doi.org/10.11896/jsjcx.220600034>

### [基于联邦学习的车联网多维资源动态分配算法](#)

Multi-dimensional Resource Dynamic Allocation Algorithm for Internet of Vehicles Based on Federated Learning

计算机科学, 2022, 49(12): 59-65. <https://doi.org/10.11896/jsjcx.211000123>

# 基于改进 NSGA-III 的多目标联邦学习进化算法

钟佳淋 吴亚辉 邓 苏 周浩浩 马武彬

国防科技大学信息系统工程重点实验室 长沙 410073

(zhongjialin20@nudt.edu.cn)

**摘要** 联邦学习技术能在一定程度上解决数据孤岛和隐私泄露的问题,但存在通信成本高、通信不稳定、参与者性能分布不均等缺点。为了改进这些缺点并实现模型有效性、公平性和通信成本的均衡,提出了一种面向联邦学习多目标优化的改进 NSGA-III 算法。首先构建联邦学习多目标优化模型,以最大化全局模型准确率、最小化全局模型准确率分布方差和通信成本为目标,提出了基于快速贪婪初始化的改进 NSGA-III 算法,提高了 NSGA-III 对于联邦学习多目标优化的效率。实验结果表明,相比经典多目标进化算法,提出的优化方法能得到较优 Pareto 解;与标准的联邦模型相比,优化的模型在保证全局模型准确率的情况下,有效降低通信成本和全局模型准确率分布方差。

**关键词** 联邦学习;多目标均衡;NSGA-III 算法;多目标进化;参数优化

**中图分类号** TP301

## Multi-objective Federated Learning Evolutionary Algorithm Based on Improved NSGA-III

ZHONG Jialin, WU Yahui, DENG Su, ZHOU Haohao and MA Wubin

Science and Technology on Information System Engineering Laboratory, National University of Defense Technology, Changsha 410073, China

**Abstract** Federated learning technology solves the problems of data islands and privacy leakage to a certain extent. However it has shortcomings such as high communication cost, unstable communication, and uneven distribution of participant performance. In order to overcome these shortcomings and achieve a balance between model effectiveness, fairness, and communication costs, an improved NSGA-III algorithm for multi-objective optimization of federated learning is proposed. First, a federated learning multi-objective optimization model is constructed to maximize the accuracy of the global model, minimize the variance of the global model accuracy distribution and minimize the communication cost of participant, and an improved NSGA-III algorithm based on fast greedy initialization is proposed, which improves the efficiency of NSGA-III for multi-objective optimization of federated learning. Experimental results show that the proposed optimization method can obtain a better Pareto solution than the classical multi-objective evolutionary algorithm. Compared with the standard model of federated learning, the optimized model can effectively lower the communication cost and the variance of the global model accuracy distribution while ensuring the accuracy of the global model.

**Keywords** Federated learning, Multi-objective equilibrium, Non-dominated sorted genetic algorithm-III (NSGA-III), Multi-objective optimization, Parameters optimization

## 1 引言

人工智能的快速发展给社会带来极大便利的同时,也带来了一些隐患,如数据孤岛和隐私泄露。传统的集中式机器学习,需要将分散的数据聚集在一起进行机器学习训练,但实际上许多领域的数据很难聚合在一起,如医院之间很难共享数据,存在严重的“数据孤岛”问题。此外,由于隐私泄露问题显现,人们隐私保护意识逐渐提高,世界各国也出台了隐私保护的法律法规,如《通用数据保护条例》(General Data Protection Regulation, GDPR)<sup>[1]</sup>规定,集中培训可能会侵犯用户隐私;《中华人民共和国网络安全法》<sup>[2]</sup>明确指出个人信息被列入重点保护范围。

因此,联邦学习<sup>[3]</sup>作为解决数据孤岛和隐私泄露问题的

可行解决方案应运而生。联邦学习可以在不聚合各参与者数据的情况下,训练出一个良好的全局模型。

但是,传统的联邦学习面临着通信成本高和结构异质性的挑战<sup>[4]</sup>。联邦学习的服务器和参与者之间的参数传输消耗着大量的通信成本;同时,不同参与者之间由于计算存储能力不同、网络环境不同,参与者在训练时会出现离线、传输模型参数丢失等情况,导致联邦学习效率、准确性、公平性等受到影响。目前有许多文献致力于通信成本或结构异质性的研究,但对这些问题的综合考虑较少。

针对以上问题,本文综合考虑了通信成本和结构异质性问题,对模型有效性、公平性和通信成本之间的平衡问题进行了研究。首先定义联邦学习为一个三目标优化模型,以同时最大化全局模型准确率、最小化全局模型准确率分布方差和

到稿日期:2022-03-04 返修日期:2022-04-22

基金项目:国家自然科学基金面上项目(61871388)

This work was supported by the General Program of National Natural Science Foundation of China(61871388).

通信作者:马武彬(wb\_ma@nudt.edu.cn)

通信成本为目标。在结合联邦学习训练特点的基础上,对第3代非支配排序遗传算法(Non-dominated Sorted Genetic Algorithm-III, NSGA-III)的初始化进行改进,设计了一种面向多目标联邦学习的基于快速贪婪初始化的 NSGA-III 算法(Fast Greedy Initialization NSGA-III, FNSGA-III)。实验结果表明,FNSGA-III 算法可以实现3个目标的均衡,能够在保持联邦学习模型整体性能不严重损失的情况下,有效降低通信成本和减小各参与者准确率分布的方差,使参与者准确率分布更均衡。本文的主要工作如下:

(1)综合考虑了全局模型准确率最大化、全局模型准确率分布方差最小化、通信成本最小化3个目标,构建了联邦学习多目标优化模型,实现了联邦学习框架下模型有效性、公平性和通信成本多个目标的均衡。

(2)提出了 FNSGA-III 算法。为了快速收敛 NSGA-III 算法并得到高质量解,提出了面向多目标联邦学习进化算法的基于快速贪婪初始化的初始解构建算法,引入二进制和实值编码和解码策略,来加快 NSGA-III 算法的进化效率。

(3)通过 MNIST 和 CIFAR-10 数据集实验,验证了 FNSGA-III 算法所得 Pareto 解优于 NSGA-III 算法,并将 FNSGA-III 算法与其他经典进化算法 NSGA-II 以及 SPEA2 进行对比,结果表明 FNSGA-III 算法所得 Pareto 解质量更高。最后,选择部分 Pareto 解,进行不同数据分布和不同联邦学习算法的增强实验,以验证所得解在不同联邦学习算法和数据分布上的扩展性。算法所得 Pareto 解能在保证联邦学习全局模型准确率的情况下,有效降低通信成本和全局模型准确率分布方差。

## 2 相关工作

近年来,联邦学习受到了广泛的关注。McMahan 等<sup>[3]</sup>于2016年首次提出了联邦学习的概念及联邦平均算法(Federated Averaging, FedAvg),该算法对隐私保护和数据孤岛下的机器学习训练有重要应用意义。联邦学习的研究不断深入,但它仍然存在一些未克服的挑战,如通信成本高、结构异质性等<sup>[4]</sup>。

为了使联邦学习在海量数据中也能适用,必须考虑降低联邦学习的通信开销的问题。McMahan 等<sup>[3]</sup>提出的 FedAvg 算法通过增加每一轮通信中的本地训练计算量,来减少全局通信轮次,从而提高通信效率。还有学者通过减小参与者上传的参数大小,来降低通信传输量。Chen 等<sup>[5]</sup>提出了一种分层异步更新算法,根据深度神经网络模型的结构特点将参数分层次为浅层参数和深层参数,在前期的全局通信迭代过程中,本地参与者与服务器之间只传输浅层参数,并只在最后几轮通信中对全局模型的深层参数进行传输聚合更新,该算法通过减少传输模型参数大小和降低神经网络中深层参数的更新频率来降低通信开销,不足之处是模型的准确率会受到影响。Zhu 等<sup>[6]</sup>将稀疏进化训练算法(Sparse Evolutionary Training, SET)<sup>[7]</sup>引入到联邦学习中,SET 算法的主要思想是在神经网络的全连接层之间,通过稀疏性参数来控制全连接网络之间的连接稀疏度,以此来减小传输模型参数的大小,有效降低了通信成本。

除了通信成本外,结构异质性也是联邦学习优化的主要问题之一。不同参与者之间由于计算存储能力不同、网络环境不同,训练时参与者会出现离线、传输模型参数丢失等情况。为了增强联邦学习的鲁棒性,有学者对结构异质性问题

进行了多方面的研究。Hao 等<sup>[8]</sup>设计了安全聚合协议,允许参与者随时退出,只要剩余参与者数能够满足联邦学习更新,就能提高系统的容错率和鲁棒性。有学者研究了如何合理分配异构设备资源,Kang 等<sup>[9]</sup>考虑了参与者的开销差异,以激励更多高质量参与者进行联邦学习训练。Li 等<sup>[10]</sup>使用全局模型准确率方差作为公平性衡量指标,设计了一个 q-FFL(q-Fair federated Learning)优化算法,该算法增加了高损耗参与者的模型聚合权重,实验结果表明,该算法能使低准确率参与者的准确率提高,实现不同参与者之间性能分布的均衡,推动联邦学习的公平资源分配。

上述研究是针对通信成本或结构异质性的某一方面进行的,在不同程度、不同目标上实现了对联邦学习算法的优化,但联邦学习的应用在更多情况下对模型的准确率、公平性、通信成本等都有要求,为了实现联邦学习框架下多个目标的平衡,有学者尝试将智能优化算法与联邦学习结合。Zhu 等<sup>[6]</sup>定义联邦学习为双目标优化问题,以最小化模型测试错误率和通信成本为目标,使用 NSGA-II(Non-dominated Sorted Genetic Algorithm-II)算法优化了联邦学习的神经网络结构参数,算法进化出的 Pareto 解与标准的 FedAvg 算法相比,在一定程度上同时提高了模型性能和通信效率,但该算法并未考虑由联邦学习的结构异质性问题带来的通信不稳定及参与者之间准确分布不均衡等其他情况,同时使用的 NSGA-II 算法对于多目标联邦学习模型的扩展性较差。Basheer 等<sup>[11]</sup>使用粒子群算法来更新优化联邦学习的神经网络的隐藏层数、神经元数和全局通信轮次,但其优化目标为单目标,未综合考虑联邦学习的其他目标。

针对上述问题,本文综合考虑通信成本和结构异质性问题,引入公平性作为优化目标,以探究联邦学习的模型准确率、公平性和通信成本的多目标均衡关系;并在实验中设定通信环境是不稳定的,增强了算法的鲁棒性。

## 3 FNSGA-III 算法

### 3.1 基本理论

#### 3.1.1 联邦学习

联邦学习是一种隐私保护机器学习技术,它使参与者共同训练出一个全局模型,而无需上传本地隐私数据到服务器中。假设有  $K$  个参与者,拥有的数据为  $\{D_1, D_2, \dots, D_K\}$ ,传统的集中式学习是把所有的数据集集中到一起,使用  $D = D_1 \cup D_2 \cup \dots \cup D_K$  来训练模型。

在联邦学习过程中,拥有数据集为  $D_k$  的第  $k$  个参与者的损失函数为:

$$L_k(\omega) = \frac{1}{n_k} \sum_{i \in D_k} l_i(\omega) \quad (1)$$

联邦学习的全局目标是 minimized 全局损失函数  $L(\omega)$ :

$$\min_{\omega} L(\omega) = \frac{\sum_{i \in \bigcup_k D_k} l_i(\omega)}{|\bigcup_k D_k|} = \sum_{k=1}^K \frac{n_k}{n} L_k(\omega) \quad (2)$$

式(1)、式(2)中, $k$ 是参与者的序号, $L_k(\omega)$ 为第  $k$  个参与者的损失函数, $l_i(\omega)$ 是数据样本  $i$  上的损失函数, $n_k$ 为参与者  $k$  的数据集  $D_k$  的大小, $n_k = |D_k|$ , $n$ 为  $K$  个参与者的数据样本总大小。联邦学习的目标是通过最小化参与者损失函数  $L_k(\omega)$  的加权平均,来优化全局损失函数  $L(\omega)$ 。联邦学习是一个协作过程,如图1所示。

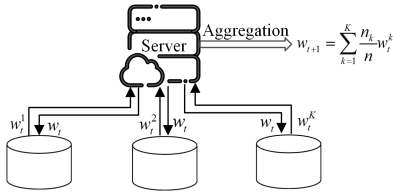


图1 联邦学习训练过程

Fig.1 Training process of federated learning

在每一轮训练中,各个参与者从服务器接收全局模型  $w_t$ ,并使用本地数据对全局模型进行训练,得到更新的本地局部模型  $w_t^k$ ,然后参与者将更新后的本地模型发送给服务器,服务器以一定的规则聚合各个模型,得到一个新的全局模型  $w_{t+1}$ ,用于下一轮次的迭代训练。下标  $t$  表示联邦学习的通信轮次。

### 3.1.2 第3代非支配排序遗传算法(NSGA-III)

基于遗传算法和 Pareto 最优解的多目标优化进化算法的研究成果有许多,如第2代非支配排序遗传算法<sup>[12]</sup>(NSGA-II)、基于分解的多目标进化算法<sup>[13]</sup>(Multi-objective Evolutionary Algorithm Based on Decomposition, MOEA/D)、SPEA2<sup>[14]</sup>(Strength Pareto Evolutionary Algorithm 2)、PAES<sup>[15]</sup>(Pareto archived evolution strategy)。NSGA-II 是一种强大的、鲁棒的多目标进化算法,适用于具有2个或3个目标的问题。如果目标数大于3,则可以采用较新的进化算法,如基于参考点的第3代非支配排序遗传算法<sup>[16]</sup>(NSGA-III),NSGA-III在4个或4个以上目标的优化问题上的性能优于NSGA-II。本文将联邦学习定义为三目标优化问题模型,为了保证算法目标的可扩展性,例如当将联邦学习的目标扩展为4个及以上时,算法仍然适用,本文采用NSGA-III算法。

### 3.2 联邦学习多目标优化模型

本文构建了联邦学习的三目标优化模型,并对其目标、决策变量、变量编码等进行了阐述,模型表达式如式(3)所示:

$$\begin{cases} \min F(\mathbf{v}) = \min (f_1(\mathbf{v}), f_2(\mathbf{v}), f_3(\mathbf{v}))^T \\ \text{s. t. } Conv \in [1, Max_{conv}], kc \in [1, Max_c], ks \in (3 \text{ or } 5) \\ L \in [1, Max_L], N \in [1, Max_N], \epsilon \in [1, Max_\epsilon] \\ \eta \in [1, Max_\eta], C \in (0, 1, 1] \end{cases} \quad (3)$$

#### (1) 目标函数

$F(\mathbf{v})$ 是模型的目标函数,此模型有3个最小化目标,分别为最小化全局模型测试错误率  $f_1$ 、最小化全局模型准确率分布方差  $f_2$ 、最小化通信成本  $f_3$ 。3个最小化函数综合考虑了联邦学习的通信成本和结构异质性问题,优化目标函数能够实现联邦学习模型有效性、公平性和通信成本之间的均衡。

传统的联邦学习没有将公平性作为优化目标,模型往往会偏向一些设备,使得最后参与者之间的准确率差距较大。将公平性目标引入联邦学习进化算法中,可以减少不同参与者之间的准确率差距,使得参与者的准确率分布更均衡,本文中公平性目标用文献<sup>[10]</sup>的全局模型准确率分布方差代表。

第三个目标函数是通信成本最小化。成本往往是越小越好,低成本意味着联邦学习训练的可行性和持续性。在联邦学习训练中,通信成本与参与者传输的模型参数直接相关,按照人工经验,很难找到模型复杂度低且准确率高的模型;同时,在神经网络中并不存在模型复杂度越高,准确率就越高的通用定律。因此,这3个目标函数形成的空间大,且3个函数

之间没有确切的相关关系。

本文中,联邦学习3个目标的具体评估过程为:对进化算法中的种群个体使用FedAvg算法,结合SET算法进行一定通信轮次的训练,再对训练好的全局模型  $w$  进行测试,得到各个参与者的准确率  $\{a_1, a_2, \dots, a_K\}$ 。计算得到全局模型平均测试准确率  $A = \frac{\sum_{k=1}^K a_k}{K}$ ,由此,可以计算出目标  $f_1$  全局模型测试错误率  $E = 1 - A$ 。

目标  $f_2$  为全局模型准确率分布方差  $V = \frac{\sum_{k=1}^K (a_k - A)^2}{K}$ 。方差可以看作公平性的度量指标之一。通过减小参与方之间全局模型准确率分布的方差,来使聚合后参与者准确率分布更均匀、公平。

目标  $f_3$  为参与者的平均通信成本,一般情况下,每个参与者的通信成本只与其传输的模型参数大小相关。由于本文使用的神经网络结构在联邦学习训练过程中并未发生改变,因此各参与者的模型参数大小  $\sigma$  相同并保持不变,则目标  $f_3$  可以表示为  $Cost = \frac{\sigma \cdot C \cdot K}{K} = \sigma \cdot C$ ,其中  $K$  是参与者总数,  $C$  是参与者每轮参与的比例,  $\sigma$  是模型参数的大小。

#### (2) 模型决策变量及其约束

因联邦学习是机器学习模型进行协作训练的过程,决策变量和约束条件为待优化的各参数及其范围,用  $\mathbf{v}$  来表示。本文设定优化的参数包括这3个部分。其中,神经网络超参数和神经网络的连通性参数  $\epsilon$  直接决定联邦学习的模型及其模型复杂度,并与每轮的参与比例  $C$  共同影响着联邦学习的通信成本、准确率和准确率分布方差3个目标函数。

本文的神经网络选择了多层感知机(Multilayer Perceptron, MLP)和卷积神经网络(Convolutional Neural Network, CNN),其中MLP模型下超参数包括隐藏层数  $L$ 、每一层隐藏层的神经元数  $N$ 、学习率  $\eta$ ; CNN超参数包括卷积层数  $Conv$ 、卷积核数目  $kc$ 、卷积核大小  $ks$ 、全连接层数  $L$ 、全连接层神经元数  $N$ 、学习率  $\eta$ 。即  $\mathbf{v} = \{Conv, kc, ks, L, N, \epsilon, \eta, C\}$ ,各个变量的取值范围在实验部分设定。

神经网络的连通性参数  $\epsilon$  借鉴了Mocanu等<sup>[7]</sup>提出的SET算法中的静态SET算法,即首先在两层全连接层之间使用ER随机图<sup>[17]</sup>初始化稀疏权重矩阵,之后保持网络的拓扑固定不变。两层之间的连接数由参数  $\epsilon$  决定,总连接数为  $n = \epsilon(n^k + n^{k-1})$ ,其中  $n^k$  和  $n^{k-1}$  分别为  $k$  层和  $k-1$  层的神经元数。本文将静态SET算法用于MLP及CNN的全连接层中。

#### (3) 决策变量编码

本文采用FNSGA-III算法来优化联邦学习的神经网络超参数、连通性参数  $\epsilon$  和每轮参与比例  $C$ 。染色体是算法操作的主体,本文中有整数和实数两种类型的决策变量,其中所有整数使用二进制编码,实数使用实值编码,进行混合编码。则MLP和CNN的卷积层数、卷积核数目、卷积核大小、全连接层数、全连接层每层神经元数和SET参数  $\epsilon$  使用二进制编码,学习率  $\eta$  和每轮参与比例  $C$  使用实值编码。MLP和CNN的编码实例如图2所示。

解码过程中,二进制解码时自动增加1,如图2(a)所示的MLP编码中,  $N_1$  编码为000111,解码为  $N_1 = 8$ 。为了方便, CNN中卷积核大小只在3~5之间选择,并始终保持卷积输出不变,则在CNN的神经网络结构中,只在最后增加一个池化层。

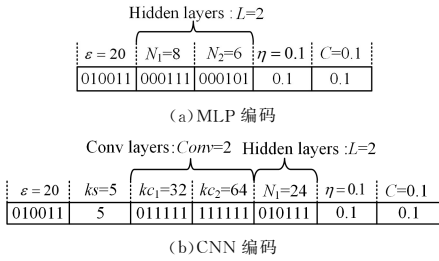


图2 MLP和CNN模型的染色体编码实例

Fig. 2 Chromosome encoding examples of MLP and CNN

### 3.3 FNSGA-III 算法的详细过程

#### 3.3.1 快速贪婪初始化

联邦学习的决策变量搜索空间大,较优的初始解能加快收敛,并且提高所得 Pareto 解的质量,得到更好的联邦学习训练参数。本文改进 NSGA-III 中的随机初始化为针对联邦学习的快速贪婪初始化,并将改进后的 NSGA-III 算法记为 FNSGA-III 算法。FNSGA-III 算法是本文的主算法,其主要思想是在 NSGA-III 初始化这一步,通过减少联邦学习的各个训练参数,并且分组并行产生初始解,来加速联邦学习的训练过程,最后选择最优结果作为初始解。快速贪婪初始化过程的简述如下:

- (1) 随机产生  $l$  倍种群初始解;
- (2) 将联邦学习的所有参与者随机分成相同大小的组,在每个组内同步进行初始解的联邦学习训练评估过程。其中联邦学习的每轮参与者数、本地训练轮次、全局通信轮次都设置为最低值,可以较快地得到联邦学习训练后的 3 个目标,并完成对所有初始解的评估;
- (3) 分别选择 3 个目标中最优的一倍种群解;
- (4) 混合去掉重复解后,从中随机选择指定种群数量解。

#### 3.3.2 FNSGA-III 算法的流程

对于联邦学习的三目标优化模型,本文使用 FNSGA-III 算法来求解,以获得一组 Pareto 最优解,其算法流程如图 3 所示。

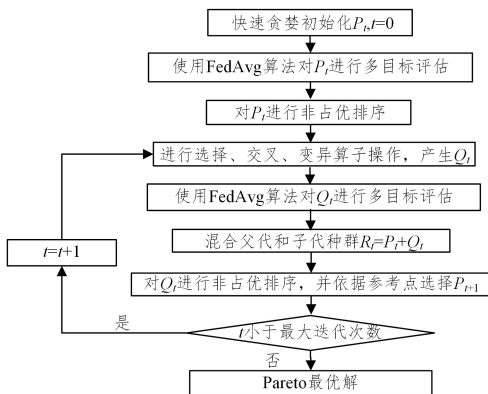


图3 FNSGA-III 算法的流程图

Fig. 3 Flow chart of FNSGA-III algorithm

FNSGA-III 首先采用快速贪婪初始化产生大小为  $N$  的初始种群,即第一代父代种群,并对变量编码。其迭代过程主要采用二元锦标赛,选择两个父代个体来产生两个子代个体,交叉变异算法分别是在二进制染色体上采用单点交叉和翻转变异,在实值染色体上采用模拟二值交叉 (Simulated Binary Crossover, SBX) 和多项式变异。重复这个过程,直到产生  $N$  个子代个体。

然后计算出子代种群的 3 个目标。混合父代种群和子代

种群,并对混合种群进行非支配排序。从中选择  $N$  个个体作为下一代的父代种群。重复这些步骤,直到满足迭代停止条件。最后得到一组 Pareto 最优解,对其进行深入分析。

其中, FNSGA-III 算法中种群个体 3 个目标的具体评估过程,是对个体进行 FedAvg 算法的训练评估。FNSGA-III 算法下的联邦学习评估过程的伪代码如算法 1 所示。

**算法 1** FNSGA-III 下种群个体的 FedAvg 三目标评估算法

输入: 种群个体  $i$  对应的决策变量参数

输出: 种群个体  $i$  的 3 个目标函数值

1. 使用  $i$  中的参数初始化稀疏神经网络权重  $w_0^i$
2. for  $t=1, 2, \dots$  do
3. Select  $m=\max(C^i \cdot K, 1)$  participants as  $S_t$
4. for each  $k \in S_t$  in parallel do
5.  $w_t^k \leftarrow$  upload from 参与者  $k$
6. end for
7. 更新全局参数  $w_{t+1} = \sum_{k=1}^K w_t^k$
8. end for
9. 使用训练后的  $w^i$  计算全局模型测试准确率  $A$  和模型参数大小  $\sigma^i = f(w^i)$
10. 计算目标  $f_1^i$  全局模型测试错误率  $E=1-A$
11. 计算目标  $f_2^i$  全局模型准确率分布方差  $V$
12. 计算目标  $f_3^i$  的通信成本  $\text{Cost}=\sigma^i \cdot C^i$
13. 参与者  $k$ :
14. download  $w_t$  as  $w^k$
15. for  $e=1, 2, \dots, E$  do
16. for  $b \in B$  do
17.  $w^k = w^k - \eta \nabla L_k(w^k, b)$
18. end for
19. end for
20. upload  $w^k$  to server

算法 1 中,  $i$  是 FNSGA-III 算法中种群的一个个体,  $P$  是种群规模。对个体  $i$  解码后,得到相关的联邦学习神经网络超参数、神经网络的连通性和参与者每轮参与比例  $C_i$ 。首先使用连通参数  $\epsilon_i$  初始化静态 SET 拓扑,将其作为算法中使用的全局模型;在每一轮训练过程中,使用小批量随机梯度下降方法 (Mini-batch SGD) 训练本地数据。经过一定轮次后,计算全局模型的测试错误率、全局模型准确率分布方差和通信成本这 3 个目标。

## 4 实验分析

### 4.1 实验设置

本节介绍了本文实验的设置,主要包括以下几个部分: 1) 实验环境及数据集; 2) 实验中使用的神经网络相关参数及连通性参数; 3) 联邦学习的参数及数据划分方式; 4) FNSGA-III 参数。

实验环境是基于 Intel(R) Core(TM) i9-9900KF CPU @ 3.60GHz  $\times$  16 的 Ubuntu 系统。在 MNIST 数据集<sup>[18]</sup> 和 CIFAR-10 数据集<sup>[19]</sup> 上进行训练测试。

本文选择 MLP 和 CNN 作为联邦学习训练的神经网络模型,并依照经验设置了本文的标准 MLP 和 CNN 参数。其中, MLP 中有 2 个隐藏层,每层有 200 个神经元(在 MNIST 数据集上参数量为 199210,在 CIFAR-10 数据集上参数量为 656810),使用 ReLU 函数作为激活函数<sup>[3]</sup>。CNN 模型设置两个  $5 \times 5$  卷积层(第一个有 32 个通道,第二个有 64 个

通道),其后接着一个  $2 \times 2$  Max 池化层和一个 128 神经元的全连接层,使用 ReLu 激活函数,最后是一个 10 类的 softmax 输出层(在 MNIST 数据集上参数量为 1 659 146,在 CIFAR-10 数据集上参数量为 2 152 266)。在 MLP 和 CNN 中,Mini-batch SGD 算法的学习率  $\eta$  为 0.05,批量大小  $B$  为 10。

在联邦学习中,本文设置的参与者总数  $K$  为 100,参与者参与比例  $C$  为 1,即每轮通信中有  $100 \times 1$  个参与者。对于参与者本地模型训练,迭代轮次 epoch 设置为 5。由于数据在不同参与者之间的大小和分布往往是不相同的,因此本文主要研究以下两个现实场景:第一种是独立同分布(Independent Identically Distribution, IID),将 MNIST 数据打乱,则 100 个参与者中每个包含 600 个样本,CIFAR-10 数据集每个参与者中包含 500 个样本;第二种是非独立同分布(Non-Independent Identically Distribution, non-IID),对 MNIST 数据集标签进行排序,将 300 个样本平均划分为 200 个片段,并给 100 个参与者分配两个片段,则每个参与者只含有两个标签且数据样本大小相同,CIFAR-10 数据集进行类似操作。同时,本文假定联邦学习的通信环境不稳定,模型参数传输存在丢失,在此设定丢失率  $Drop=30\%$ ,在通信过程中进行随机丢失。

FNSGA-III 参数设置中,各个算子的设置采用文献[6]的相关设置,选择算子采用二轮锦标赛,二进制染色体中采用概率为 0.9 的单点交叉和概率为 0.1 的位翻转变异;实值染色体中采用概率为 0.9 且  $n_c=2$  的模拟二进制交叉,和概率为 0.1 且  $n_m=20$  的多项式变异。

## 4.2 多目标联邦学习进化算法分析

### 4.2.1 FNSGA-III 的性能分析

首先在 MNIST 和 CIFAR-10 数据集上对 FNSGA-III 算法和 NSGA-III 算法进行对比实验,探究提出的快速贪婪初始化方法的有效性。实验的相关参数设置如表 1 所列。

表 1 多目标联邦学习进化算法相关参数设置

Table 1 Related parameter settings of multi-objective federated learning evolutionary algorithm

参数	MLP	CNN
种群大小	20	20
最大迭代次数	20	20
联邦学习参与比例	0.1~1	0.1~1
学习率	0.01~0.2	0.01~0.2
SET 参数	1~128	1~128
MLP 隐藏层数	1~4	
MLP 隐藏层神经元	1~256	
CNN 卷积层数		1~3
CNN 卷积核通道数		1~64
CNN 卷积核大小		3 or 5
CNN 全连接层数		1~3
CNN 全连接层神经元		1~256

本文设置种群大小为 20,种群迭代代数 20,对每个个体进行联邦学习评估过程的通信轮数为 5。设置参与者参与比例  $C$  的取值范围为 0.1~1,以保证有一定的参与者参与训练,学习率为 0.01~0.2,太高的学习率会影响收敛。

神经网络的参数设置中,MLP 的最大隐藏层数为 4,每层的最大神经元数为 256。对于 CNN,设置最大卷积层数为 3,最大核通道数为 64,最大全连接层数为 3,卷积层中最大神经元数为 256,卷积核大小为 3 或 5。网络稀疏度参数最大值设置为 128。

在 MNIST 和 CIFAR 数据集的 IID 和 non-IID 两种数据分布上,对 MLP 和 CNN 模型使用 FNSGA-III 算法和 NSGA-III 算法,进化得到的最终 Pareto 解分布如图 4 和图 5 所示,其中每个点代表联邦学习中特定结构参数对应的一个解。红色的点表示采用基于快速贪婪初始化的 FNSGA-III 算法得到的 Pareto 最优解,蓝色的点表示采用随机初始化的 NSGA-III 算法得到的 Pareto 最优解。

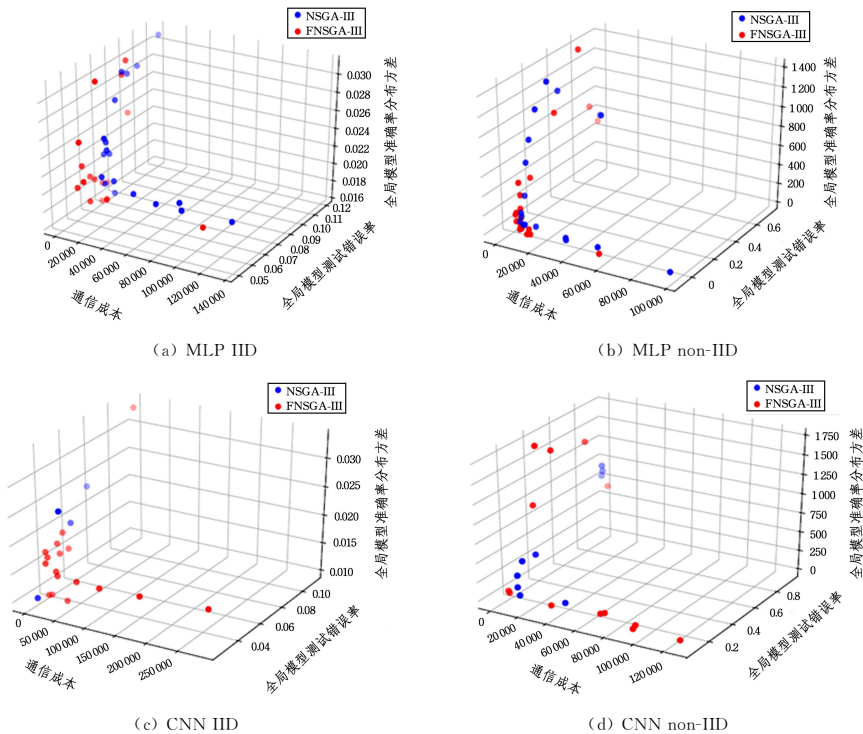


图 4 MNIST 中 FNSGA-III 与 NSGA-III 算法的 Pareto 解(电子版为彩图)

Fig. 4 Pareto solutions of FNSGA-III and NSGA-III in MNIST

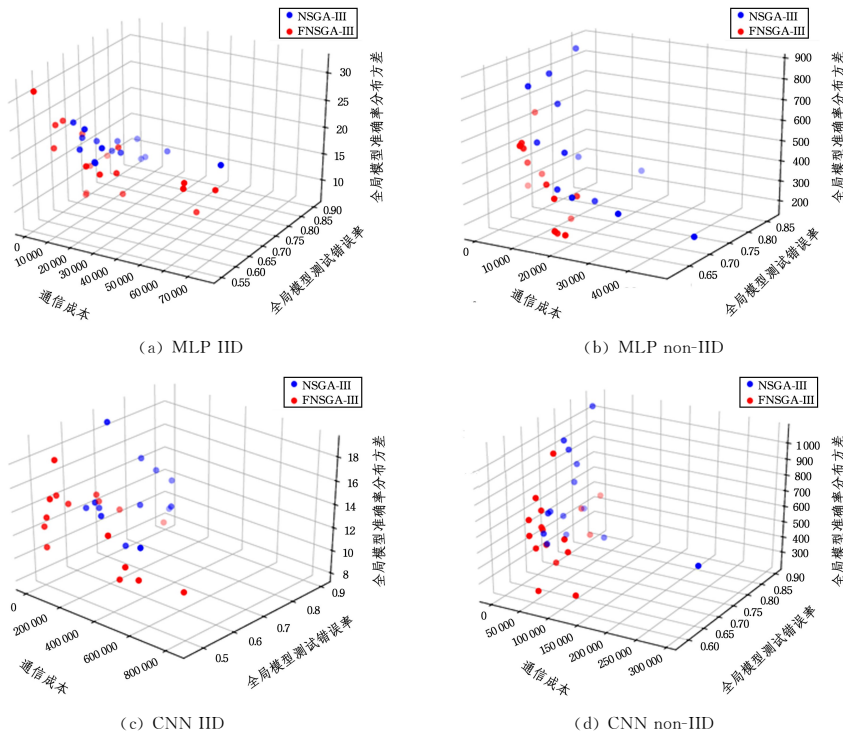


图5 CIFAR-10中FNSGA-III与NSGA-III算法的Pareto解(电子版为彩图)

Fig. 5 Pareto solutions of FNSGA-III and NSGA-III in CIFAR-10

从图4和图5可以看出,红色解基本占优于蓝色解,除了图4(c)MNIST数据集中CNN IID下一个蓝色解占优于红色解,但此CNN IID的解数量非常少,即FNSGA-III算法得到Pareto解基本优于NSGA-III算法的Pareto解。同时,可以发现,FNSGA-III更收敛于拐点处,拐点处的解特点是各个

目标值小、解质量更高。说明MNIST数据集的拐点解相对CIFAR-10数据集更集中。

此外,表2列出了FNSGA-III与NSGA-III最后得到的Pareto解的相关评价指标结果,从多个维度对两种算法的Pareto解进行了分析。

表2 FNSGA-III算法与NSGA-III算法的各指标分析

Table 2 Indicators comparison of FNSGA-III algorithm and NSGA-III algorithm

		MNIST数据集				CIFAR-10数据集			
		MLP		CNN		MLP		CNN	
		IID	non-IID	IID	non-IID	IID	non-IID	IID	non-IID
超体积 HV	FNSGA-III	1735.80	160217675	740.55	419212793	493196	6825789	2919025	67680236
	NSGA-III	1360.80	152034760	738.12	391548423	336662	4743992	2047541	42010031
Pareto 最优解数	FNSGA-III	16	20	17	13	18	15	16	17
	NSGA-III	19	18	4	9	16	13	13	14
覆盖率/%	$C(F,N)$	79	72	50	56	68.75	92.30	76.92	71.40
	$C(N,F)$	0	0	53	8	5.50	0	0	5.80
参与者平均运行时间/s	FNSGA-III	1797	4575	2811	5947	1868	1848	2679	2454
	NSGA-III	2429	4980	3305	6698	2356	2054	3653	2669
最小全局模型测试错误率/%	FNSGA-III	4.94	11.02	2.52	7.22	54.92	61.49	46.89	58.22
	NSGA-III	5.15	11.81	2.45	12.11	58.75	62.11	49.76	67.03
最小方差	FNSGA-III	2.76	25.59	1.04	15.45	7.50	179.10	7.99	242.50
	NSGA-III	3.50	25.96	1.15	83.88	7.60	209.50	9.20	339.50
最小通信成本	FNSGA-III	704	226	803	2784	1222	529	1789	755
	NSGA-III	1909	302	4744	1556	2481	1141	7589	1572

综合性指标超体积指标<sup>[20]</sup>(Hypervolume, HV)计算的是由所有非支配解与参考点构成的超立方体的超体积之和,一般来说, HV值越大,所评估的Pareto解的质量越好。由表2可知,FNSGA-III算法的HV值始终优于NSGA-III,表现出了更好的质量。

FNSGA-III算法获得的Pareto非支配解数量稳定,除了MNIST数据集的MLP IID以外,其余数量都大于NSGA-III。FNSGA-III算法在解数量上比NSGA-III算法

更具有鲁棒性。

覆盖率 $C(A,B)$ <sup>[21]</sup>计算的是解集B中的解至少被A中的一个解支配的比例,衡量的是两个解集之间的重合程度,指标C越大,代表解集A的质量较解集B的质量越好。表2中, $C(F,N)$ 中的F表示FNSGA-III的解集,N表示NSGA-III解集, $C(F,N)$ 度量值基本都大于 $C(N,F)$ ,MNIST数据集中CNN IID下的 $C(F,N) = 50\%$ 小于 $C(N,F) = 53\%$ ,但两者差距很小。从覆盖率C来看,FNSGA-III的

解优于 NSGA-III 算法。

运行时间上,实验是在一台主机上进行联邦学习的仿真实验,运行总时间是所有参与者运行时间的累加和,则参与者的平均运行时间更能表示进化算法的实际运算时间。因本文设置 FNSGA-III 进化算法中参与者参与比例  $C$  的取值范围为  $0.1 \sim 1$ ,得到参与者平均运行时间=运行总时间/平均参与者数量,如表 2 所列。可以发现,在两个数据集中,由于 CNN 模型较 MLP 复杂,CNN 时间消耗都大于 MLP,且 NSGA-III 时间的消耗大于 FNSGA-III,即从统计结果来看,FNSGA-III 的时间性能较优。

单一目标的最小值反映了各个目标函数的极值情况,体现了算法的寻优能力。从表 2 可以看出,FNSGA-III 得到的

单一目标下的最小值基本都比 NSGA-III 小,寻优能力更强。其中因为数据分布不同,non-IID 的错误率和方差都大于 IID。

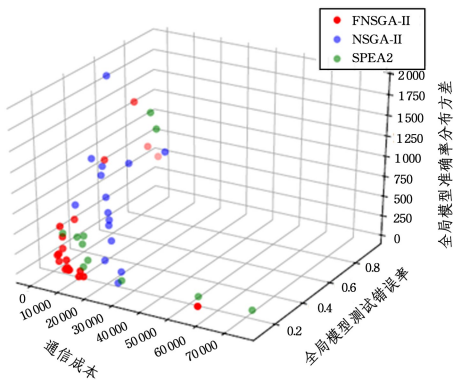
根据以上对 MNIST 和 CIFAR-10 数据集进行分析,可以得出结论,提出的 FNSGA-III 算法得到的 Pareto 最优解具有更好的质量,优于 NSGA-III 算法。

此外,本文将提出的算法 FNSGA-III 与 NSGA-II, SPEA2 进化多目标算法进行对比,同样从 HV 综合指标、Pareto 解数目、覆盖率  $C$ 、时间、单目标最优解等多方面来评估算法所得 Pareto 最优解的优劣。选择上述实验中,FNSGA-III 与 NSGA-III 所得 Pareto 解中错误率最高的 MLP non-IID,在 MNIST 和 CIFAR-10 数据集中进行多个进化算法的对比实验。实验结果如表 3 和图 6 所示。

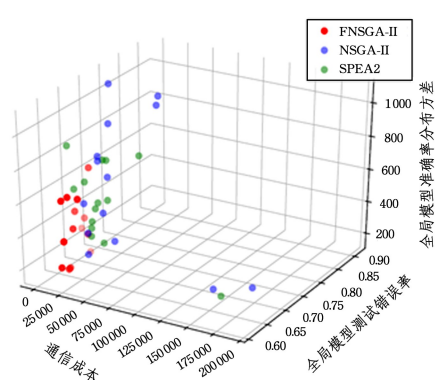
表 3 FNSGA-III 算法与其他进化算法在 MNIST 和 CIFAR-10 数据集上进行 MLP non-IID 实验的各指标分析

Table 3 Indicators comparison of FNSGA-III and other evolutionary algorithms in MLP non-IID experiment on MNIST and CIFAR-10 datasets. [F]

	MNIST			CIFAR-10		
	FNSGA-III	NSGA-II	SPEA2	FNSGA-III	NSGA-II	SPEA2
超体积量 HV	174922748	155347122	160855987	48858443	40837531	36304737
Pareto 最优解数	20	15	11	15	14	14
覆盖率/%	—	100	91	—	92.85	85.7
参与者平均运行时间/s	4575	8622	4366	1848	2485	2326
最小全局模型测试错误率/%	11.02	12.49	13.58	61.17	61.49	65.20
最小方差	25.59	66.73	76.58	169.10	202.60	179.10
最小通信成本	226	1672	368	529	2448	4247



(a) MNIST, MLP non-IID



(b) CIFAR-10, MLP non-IID

图 6 FNSGA-III, NSGA-II 和 SPEA2 的 Pareto 最优解(电子版为彩图)

Fig. 6 Pareto-optimal solutions of FNSGA-III, NSGA-II and SPEA2

对表 3 和图 6 进行简单分析,表 3 中,FNSGA-III 的 HV 值、Pareto 解数量、覆盖率、三目标下的最小值基本都优于 NSGA-II 和 SPEA2 的相关指标。时间上,MNIST 数据集中,SPEA2 的运行时间最短,但 FNSGA-III 的运行时间与 SPEA2 相差不大。由图 6 可以发现,FNSGA-III 的 Pareto 解更集中在拐点,并且占优于 NSGA-II 和 SPEA2 的 Pareto 解,即 3 个目标都优于 NSGA-II 和 SPEA2 算法。

综上,提出的 FNSGA-III 算法将随机初始化改为快速贪婪初始化,能够在提高运行效率的同时基本优于 NSGA-III, NSGA-II 以及 SPEA2 进化算法,得到的 Pareto 质量更高。

#### 4.2.2 对 FNSGA-III 所得 Pareto 解进行增强实验

由于在 FNSGA-III 的联邦学习评估过程中的通信轮次设置非常小,因此未能充分探究其联邦学习性能。由于计算资源的限制,在增强实验部分,本文只选择准确率最差的 MLP non-IID 进行增强实验,以此验证 FNSGA-III 算法所得解优于经典进化算法所得解和本文设定的联邦学习标准解,从而实现联邦学习多目标均衡及超参数寻优,同时验证所得解在不同数据分布和不同算法下的实验效果,探究所得解的可扩展性。

在两个数据集中,分别对 FNSGA-III 算法得到的 MLP non-IID 的 Pareto 最优解选择两个解,其中 1 个选择全局测试错误率很小的解记为解 1,另外 1 个选择拐点解记为解 2。

并分别选择进化算法 NSGA-II 和 SPEA2 的错误率最小解, 将其分别记为 NSGA-II 解和 SPEA2 解。本文进化算法的联邦评估过程是在 FedAvg 算法上进行的。将选择解和标准解在 FedAvg 和 FedProx 算法<sup>[22]</sup>上进行实验, FedProx 算法是 FedAvg 的改进算法, 主要面对的是 non-IID 数据, 以验证所得解对应的神经网络模型在不同的联邦学习算法中具有可迁移性, 即所得解具有良好的扩展性。

本文将通信轮次设置为 150 轮。除了增加通信轮次, 各个解在 IID 和 non-IID 下进行验证, 研究在 non-IID 数据集上

得到的非支配解是否在 IID 数据集上仍然有效。所有验证结果如表 4 和表 5 所列, 全局测试精度如图 7 和图 8 所示。

对所选择的 MLP non-IID 下的 Pareto 解的演化情况进行分析。从表 4 和表 5 可以发现, FNSGA-III 算法的解 1 都有最高的准确率, 但方差不一定最小, 体现了多目标均衡。当准确率最高时, 其通信成本或者方差可能有所损失。同时, 各选择解的通信成本明显比标准解低, 客户端参与比例也较小, 实现了从客户端参与比例和模型复杂度两方面降低通信成本。

表 4 MNIST 数据集中 FNSGA-III 算法所得 MLP non-IID 解的对比实验数据

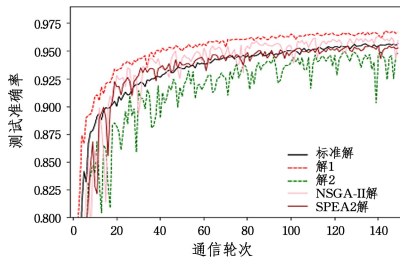
Table 4 Comparative experiment data of MLP non-IID solutions obtained by FNSGA-III in MNIST dataset

Parameter	解 1	解 2	NSGA-II 解	SPEA2 解	标准解	
参与比例 $C$	0.6200	0.2812	0.2856	0.4688	1.0000	
学习率 $\eta$	0.1125	0.0850	0.0949	0.1665	0.0500	
SET 参数 $\epsilon$	95	31	48	17	—	
MLP 隐藏层神经元数	173	197	[158, 118]	[195, 214]	[200, 200]	
通信成本	57553	9163	17116	12260	199210	
FedAvg, non-IID	准确率/%	96.70	94.81	95.96	95.28	95.69
	方差	4.71	12.78	10.50	12.36	9.86
FedAvg, IID	准确率/%	97.98	97.45	97.77	97.52	97.35
	方差	1.67	2.18	1.97	2.43	2.24
FedProx, non-IID	准确率/%	95.90	93.97	95.49	93.92	93.46
	方差	6.82	15.05	12.18	16.61	14.60
FedProx, IID	准确率/%	97.65	97.21	97.54	96.84	87.38
	方差	2.05	2.40	2.15	2.57	8.73

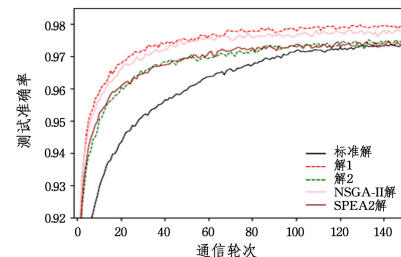
表 5 CIFAR-10 数据集中 FNSGA-III 算法所得 MLP non-IID 解的对比实验数据

Table 5 Comparative experiment data of MLP non-IID solutions obtained by FNSGA-III in CIFAR-10 dataset

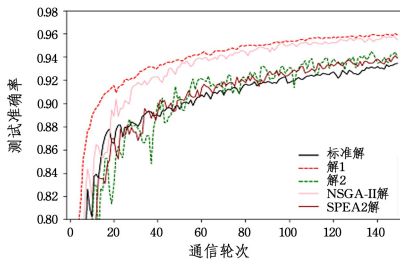
Parameter	解 1	解 2	NSGA-II 解	SPEA2 解	标准解	
参与比例 $C$	0.2940	0.2671	0.2750	0.4700	1.0000	
学习率 $\eta$	0.0931	0.1110	0.0955	0.0496	0.0500	
SET 参数 $\epsilon$	88	22	39	48	—	
MLP 隐藏层神经元数	[235]	[134]	[91]	[109, 60]	[200, 200]	
通信成本	293611	72015	124368	160007	656810	
FedAvg, non-IID	准确率/%	46.01	45.6	35.66	43.99	44.48
	方差	349.8	339.3	438.2	119.02	213.2
FedAvg, IID	准确率/%	53.33	50.02	48.83	50.62	46.79
	方差	21.88	18.71	20.47	21.97	21.84
FedProx, non-IID	准确率/%	48.53	43.34	44.45	45.71	43.09
	方差	323.87	211.98	158.30	96.56	174.06
FedProx, IID	准确率/%	52.56	50.83	49.70	49.94	45.53
	方差	33.40	29.10	25.030	25.30	19.20



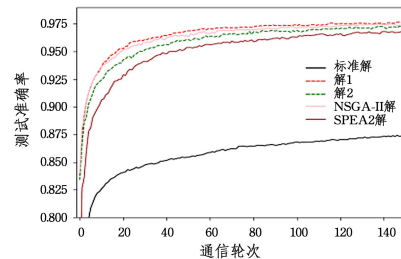
(a) FedAvg, non-IID



(b) FedAvg, IID



(c) FedProx, non-IID



(d) FedProx, IID

图7 MNIST 数据集中 FNSGA-III 算法所得的 MLP non-IID 解迭代曲线

Fig. 7 Iterative curve of MLP non-IID solutions obtained by FNSGA-III in MNIST dataset

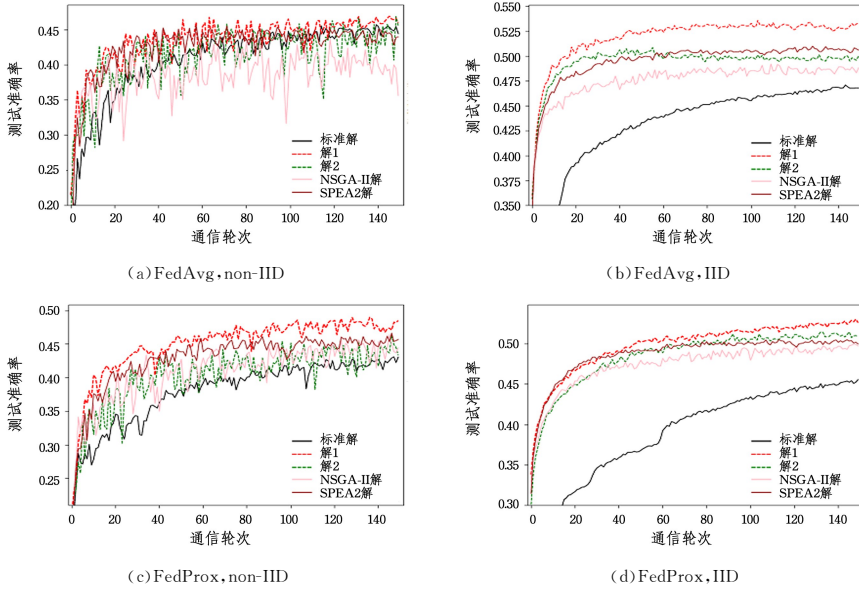


图8 CIFAR-10 数据集中 FNSGA-III 算法所得的 MLP non-IID 解迭代曲线

Fig. 8 Iterative curve of MLP non-IID solutions obtained by FNSGA-III in CIFAR-10 dataset

图7和图8中,FNSGA-III算法的解1在不同数据集、不同数据分布下都有最高准确率且迭代曲线平稳。同时可以发现,NSGA-II解和SPEA2解迭代曲线效果比标准解好,但仍劣于FNSGA-III的解1,验证了FNSGA-III算法对比经典进化算法能得到更高质量的解。

在数据分布和算法扩展上,两个数据集的所有 MLP non-IID解扩展到不同算法的 IID下,都能优于标准解,体现算法所得解在数据分布上的良好扩展性。并且,从 FedAvg 算法扩展运用到 FedProx 算法时,所有选择解均能优于标准解,验证了所得解在不同算法上的扩展性。

在 CIFAR-10 的 FedAvg, non-IID 中,出现了各解波动明显的情况,但此情况下,解1迭代曲线相对平稳,其他解也在标准解上下波动,说明各选择解仍有效。CIFAR-10 相比 MNIST 数据集的 non-IID 分布的数据倾斜程度更高,因此各解迭代曲线波动明显,将解放在针对 non-IID 数据的 FedProx 算法下,则迭代平稳。

通过分析可知,FNSGA-III 算法所得 MLP non-IID 下的解在 non-IID 下有效,扩展到 IID 下仍能有较好的运行效果。并且,所得解从 FedAvg 算法扩展到 FedProx 算法中仍有较好的表现,体现了所得解在数据分布和算法上都有较好的扩展性。并且,两个数据集的选择解在准确率上基本优于标准解和经典算法 NSGA-II 和 SPEA2 的解,实现了对联邦学习的多目标均衡和超参数寻优。

**结束语** 本文提出 FNSGA-III 算法来解决多目标联邦学习模型问题,并在通信不稳定的情况下进行实验验证。首先构建了联邦学习的三目标模板,设定优化目标为最小化全局模型测试错误率、通信成本、全局模型准确率分布方差,决策变量为神经网络的超参数以及联邦学习参数。引入 NS-

GA-III 算法对联邦学习多目标模型进行求解,并改变了 NSGA-III 的初始化,实验结果表明,改进的 FNSGA-III 算法优于原 NSGA-III 算法。并且使用 FNSGA-III 算法优化得到的 Pareto 最优解在不同数据分布和不同算法上具有可扩展性,与基准的联邦学习模型相比,有效提高了全局模型准确率,降低了全局模型准确率分布方差及通信成本,实现了对多目标联邦学习的超参数寻优。

本文实现的多目标联邦学习进化算法能够实现不同目标的均衡,但仍存在计算时间长的问题,如何进一步提高多目标联邦学习的计算效率是未来的研究方向之一。

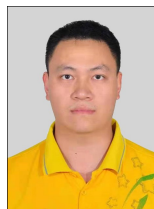
## 参考文献

- [1] REGULATION G D P. Regulation EU 2016/679 of the European Parliament and of the Council of 27 April 2016[J]. Official Journal of the European Union, 2016(59): 1-88.
- [2] SUN Y H. Cybersecurity Law: The Fundamental Measures to Ensure Cybersecurity—Study and implement the “Network Security Law of the People’s Republic of China”[J]. China Information Security, 2016(12): 30-33.
- [3] MCMAHAN B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data [C]// Artificial Intelligence and Statistics. PMLR, 2017: 1273-1282.
- [4] LI L, FAN Y, TSE M, et al. A review of applications in federated learning [J]. Computers & Industrial Engineering, 2020, 149(5): 106854.
- [5] CHEN Y, SUN X, JIN Y. Communication-efficient federated deep learning with layerwise asynchronous model update and temporally weighted aggregation [J]. IEEE Transactions on

- Neural Networks and Learning Systems, 2019, 31(10): 4229-4238.
- [6] ZHU H, JIN Y. Multi-objective evolutionary federated learning [J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 31(4): 1310-1322.
- [7] MOCANU D C, MOCANU E, STONE P, et al. Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science[J]. Nature Communications, 2018, 9(1): 1-12.
- [8] HAO M, LI H, LUO X, et al. Efficient and privacy-enhanced federated learning for industrial artificial intelligence[J]. IEEE Transactions on Industrial Informatics, 2019, 16(10): 6532-6542.
- [9] KANG J, XIONG Z, NIYATO D, et al. Incentive design for efficient federated learning in mobile networks: A contract theory approach[C]// 2019 IEEE VTS Asia Pacific Wireless Communications Symposium. IEEE, 2019: 1-5.
- [10] LI T, SANJABI M, BEIRAMI A, et al. Fair resource allocation in federated learning[J]. arXiv:1905.10497, 2019.
- [11] QOLOMANY B, AHMAD K, AL-FUQAHA A, et al. Particle swarm optimized federated learning for industrial IoT and smart city services[C]// 2020 IEEE Global Communications Conference(GLOBECOM 2020). IEEE, 2020: 1-6.
- [12] DEB K, PRATAP A, AGARWAL S, et al. A fast and elitist multi-objective genetic algorithm: NSGA-II [J]. IEEE Transactions on Evolutionary Computation, 2002, 6(2): 182-197.
- [13] ZHANG Q, LI H. MOEA/D: A multi-objective evolutionary algorithm based on decomposition[J]. IEEE Transactions on evolutionary computation, 2007, 11(6): 712-731.
- [14] ZITZLER E, LAUMANN S, THIELE L. SPEA2: Improving the strength Pareto evolutionary algorithm: TIK-Report 103 [R]. Swiss/Swiss Federal Institute of Technology, 2001.
- [15] KNOWLES J, CORNE D. The pareto archived evolution strategy: A new baseline algorithm for pareto multi-objective optimization[C]// Proceedings of the 1999 Congress on Evolutionary Computation-CEC99. IEEE, 1999: 98-105.
- [16] DEB K, JAIN H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints[J]. IEEE Transactions on Evolutionary Computation, 2013, 18(4): 577-601.
- [17] ERDOS P, RÉNYI A. On the evolution of random graphs[J]. Publications of the Mathematical Institute of the Hungarian Academy of Sciences, 1960, 5(1): 17-60.
- [18] LECUN Y, CORTES C, BURGESS J. MNIST handwritten digit database[J/OL]. <http://yann.lecun.com/exdb/mnist>.
- [19] KRIZHEVSKY A, HINTON G. Learning multiple layers of features from tiny images[J/OL]. <http://www.cs.toronto.edu/~kriz/cifar.html>.
- [20] ZITZLER E, THIELE L. Multi-objective evolutionary algorithms: a comparative case study and the strength Pareto approach[J]. IEEE Transactions on Evolutionary Computation, 1999, 3(4): 257-271.
- [21] ZITZLER E, THIELE L. Multi-objective optimization using evolutionary algorithms—a comparative case study[C]// International Conference on Parallel Problem Solving From Nature. Berlin: Springer, 1998: 292-301.
- [22] LI T, SAHU A K, ZAHEER M, et al. Federated optimization in heterogeneous networks[C]// Proceedings of Machine Learning and Systems. 2020: 429-450.



**ZHONG Jialin**, born in 1998, postgraduate. Her main research interests include federated learning and so on.



**MA Wubin**, born in 1986, Ph.D, associate researcher. His main research interests include multi-objective optimization, micro service and data mining.

(责任编辑:喻黎)