文章编号:1671-4598(2025)11-0252-07 DOI:10.16526/j.cnki.11-4762/tp.2025.11.030 中图分类号:TP183 文献标识码:A

基于申威服务器的联邦边缘学习数据队列优化研究

王晓博，谷洪峰，陈盛龙
（无锡先进技术研究院，江苏 无锡 214000）

摘要：物联网中的联邦边缘学习系统被认为是一种有前景的技术，可以在保证客户端隐私的同时减少计算负担；在联邦边缘学习系统中，联邦边缘的数据队列是有限的，且联邦边缘与客户端之间的信道是时变的；因此，选择合适数量的客户端上传数据以保持数据队列的稳定，同时最大化学习精确度，是一项挑战；为了解决这一问题，通过结合联邦边缘学习系统和信息瓶颈理论，对如何高效使用数据队列进行了研究，提出了一种基于申威服务器的数据队列优化方法；具体来说，采用 Lyapunov 优化理论确定最优选择客户端数，在学习精确度和队列稳定性之间实现平衡；进一步使用信息瓶颈理论在保持数据队列大小不变的情况下，最大化数据的压缩和存储效率；通过仿真实验对该方法的性能进行了评估，结果表明，所提出的方法优于已知的基准方法，提升了数据的存储和处理能力，为物联网系统的设计与优化提供了新的思路。

关键词：联邦边缘学习；信息瓶颈；数据队列优化；稳定性；精度

Research on Optimization of Federated Edge Learning Data Queue Based on Sunway Server

WANG Xiaobo, GU Hongfeng, CHEN Shenglong
(Wuxi Institute of Advanced Technology, Wuxi 214000, China)

Abstract: The federated edge learning (FEEL) system in the internet of things (IoT) is regarded as a promising technology that can reduce computational complexity while ensuring client privacy. In this system, the data queue at the federated edge is limited, and the channel between the federated edge and clients is time-varying. Therefore, it is a significant challenge for selecting an appropriate number of clients to upload data to maintain the stability of the data queue while maximizing learning accuracy To address this issue, by combined the federated edge learning system with the information bottleneck theory, research on the efficient utilization of the data queue is conducted, and a data queue optimization method based on the Sunway server is proposed. Specifically, The Lyapunov optimization theory is employed to determine the optimal number of selected clients, so as to achieve a balance between the learning accuracy and queue stability. Furthermore, the information bottleneck theory is applied to maximize the data compression and storage efficiency with the data queue size unchanged. Simulation experiments are conducted to evaluate the performance of the proposed method, and the results demonstrate that the method outperforms the established benchmark approaches, enhancing data storage and processing capabilities and providing new insights for the design and optimization of IoT systems.

Keywords: federated edge learning; information bottleneck; data queue optimization; stability; accuracy

0 引言

随着物联网和社交网络应用的快速发展，并且很多设备配备了各种传感器和日益强大的硬件，导致设备生成的数据呈指数增长。据预测，数据生成速度将在不久的将来超过现今互联网的容量[1-2]。由于网络带宽和数据隐私的考虑，将所有数据发送到远程云端是不切实际的。因此，研究机构估计超过 90% 的数据将会在本地存储和处理，从而满足如智能家居、健康监测、智慧城市、智能交通和智能安防等设备应用的需求[3-6]。然而，设备数据往往是私密的，不同设备不愿意与其它设备分享自己的数据，因此这些应用因为缺乏训练数据而难以生成精确模型。为了解决这个问题，引入了联邦学习，通过共享模型而非共享原始数据来保护设备隐私[7-8]。

收稿日期:2024-10-17； **修回日期**:2024-11-27。
作者简介:王晓博(1998-),男,硕士。
引用格式:王晓博,谷洪峰,陈盛龙.基于申威服务器的联邦边缘学习数据队列优化研究[J].计算机测量与控制,2025,33(11):252-258.

具体来说，每个客户端通过本地训练数据获取一个本地模型，然后将该本地模型上传到中央云服务器。中央云服务器聚合本地模型获得全局模型，然后将全局模型发送给客户端，从而满足相关应用[9]。通过这种方式，客户端可以在保护数据的同时获得精确的模型。然而，联邦学习中的本地训练可能会对计算能力有限的客户端造成较大的计算负担[10-12]。

在本文中，考虑了一种包括中央云服务器、联邦边缘和客户端的三层基础设施，即联邦边缘学习系统。它可以减轻客户端的计算负担，并保护客户端隐私。具体来说，联邦边缘收集并训练客户端的数据以获得本地模型，然后将结果返回给客户端，并删除数据。随后，中央云服务器收集和聚合来自联邦边缘的所有本地模型以获得全局模型。最后，联邦边缘从中央云服务器下载全局模型，为之后的训练做参考[13-14]。

在联邦边缘学习系统中，客户端上传用于训练的数据会暂时存储在每个联邦边缘的数据队列中。每个联邦边缘的数据队列是有限的。如果过多客户端向其上传数据，数据队列就会因为溢出而变得不稳定。如果上传数据的客户端太少，数据队列中的数据不足，模型的精确性就会下降。为了在保证队列稳定的同时获得精确的模型，客户端数据应该尽量填满联邦边缘的数据队列。另一方面，训练会在数据到达数据队列后立即开始。一旦训练完成，训练结果会发送回客户端，并且只有在信道条件足够好的情况下，使用过的数据才会从数据队列中删除。选择合适数量的客户端上传数据到数据队列是具有挑战性的。

另外信息瓶颈作为一种新颖的数据压缩和特征提取方法，在物联网系统中的应用日益引起人们的关注。信息瓶颈理论通过利用互信息来压缩数据以发现数据中的关键信息，在保留关键信息的同时，压缩和提取最相关的特征，从而实现更高效的数据压缩和存储。这为解决数据队列优化问题带来了一个新的可能性[15-16]。

因此，结合联邦边缘学习系统和信息瓶颈理论的研究，有助于数据队列的优化，提升系统的通信效率和数据处理能力。这一研究方向具有重要的理论意义和实践价值，将为物联网系统的设计和优化提供新的思路，推动数据队列优化进一步发展和应用。本文的主要贡献如下：

1）在物联网中考虑了联邦边缘学习系统，并提出了一种数据队列优化方法，以最大化训练模型的精确性，同时确保数据队列的稳定性。

2）对于每个联邦边缘，基于 Lyapunov 优化理论，制定了优化问题，以确定能够在保持数据队列稳定的同时获得高模型精度的客户端数量。

3）结合信息瓶颈理论，通过利用互信息，在保留关键信息的同时，压缩和提取最相关的特征，从而实现更高效的数据压缩和存储。

同时，回顾了一些关于联邦学习、联邦边缘学习、数据队列优化的相关研究工作。

文献［17］为了克服隐私问题和较高的通信延迟，提出了一种基于中央处理器或图形处理器的联邦学习系统来加快训练过程，从而提高学习效率。文献［18］考虑了一种具有模拟梯度聚合的联邦学习系统，提出了一种能量感知的动态设备调度算法，从而提高精度。文献［19］利用传输功率控制来减少联邦学习系统的聚合误差，考虑了一种新的功率控制设计，从而最小化均方误差。文献［20］研究了联邦学习空中模型聚合，开发了一种基于消息传递的算法，以降低复杂度和提高性能。然而，上述联邦学习的相关研究工作，都是客户端在本地训练模型，这会增加客户端的通信开销和计算负担。

文献［9］研究了一种三层联邦边缘学习系统，提出了一种高效的资源调度算法，从而实现消耗资源最小化。文献［21］为了解决局部模型过度拟合局部数据的问题，设计了一个在边缘服务器执行联邦梯度下降，在云层执行联邦平均的三层联邦边缘学习系统，从而增强了联邦平均性能。文献［22］提出了一个基于车联网的联邦边缘学习客户端选择框架，用于解决数据和系统异构性下复杂的客户端选择问题。然而，虽然上述联邦边缘学习的相关研究工作中客户端是通过服务器的帮助进行模型训练，以减轻用户的通信开销和计算负担，但是它们都没有考虑数据队列的优化。

文献［23］考虑了海量数据进行大数据实时处理时，需要一个性能优良的分布式消息队列的问题，通过分析 3 种分布式消息队列，分别进行实时计算实验。文献［24］考虑了不同差异性会给分布式系统应用间的通信带来很大麻烦的问题，提出了 RabbitMQ 在配置信息分发业务中的一种使用模型。文献［25］提出了一种基于排队论的实时以太网缓存队列优化算法，确定了数据帧排队延时是影响网络延时的主要因素。然而，虽然上述研究工作考虑了数据队列的优化，但是它们都没有考虑使用变分信息瓶颈方法对数据进行高效压缩。这促使我们开展这项研究工作。

1　联邦边缘学习系统

提出的联邦边缘学习系统如图 1 所示，由客户端、联邦边缘和中央云服务器组成。每个联邦边缘配备一个数据队列，队列遵循先到先服务（FCFS，first come first serve）策略，最大长度为 $Q_{\max}$。客户端包括智能手机、平板电脑和笔记本电脑等，它们通常具有有限的计算资源。与之对应，联邦边缘具备能够完全满足客户端各类应用的计算资源。假设每个联邦边缘覆盖范围内

有 k 个客户端。客户端向对应联邦边缘上传数据的时间被划分为相等的时隙，每个时隙中，客户端会消耗部分电量上传数据。首先，联邦边缘根据当前数据队列的积压情况以及每个客户端的剩余电量，选择适量的客户端上传数据。在确保队列不溢出的前提下，尽可能让更多的数据参与训练。在数据到达联邦边缘的数据队列后，训练立即开始，并在训练结束后联邦边缘更新本地模型。当联邦边缘和一些客户端之间的通信条件足够好时，联邦边缘将训练结果发回给这些客户端，并从数据队列中删除已使用的训练数据。否则，已使用的数据将继续积压在数据队列中，直到通信条件变得足够好。然后，进入下一时隙，重复上述过程，更新本地模型。当还有剩余电量的客户端没有数据可上传时，联邦边缘将最终更新的本地模型上传到中央云服务器。当所有联邦边缘都上传了本地模型后，中央云服务器聚合所有本地模型为一个全局模型。最后，联邦边缘从中央云服务器下载全局模型，为之后的训练做参考。

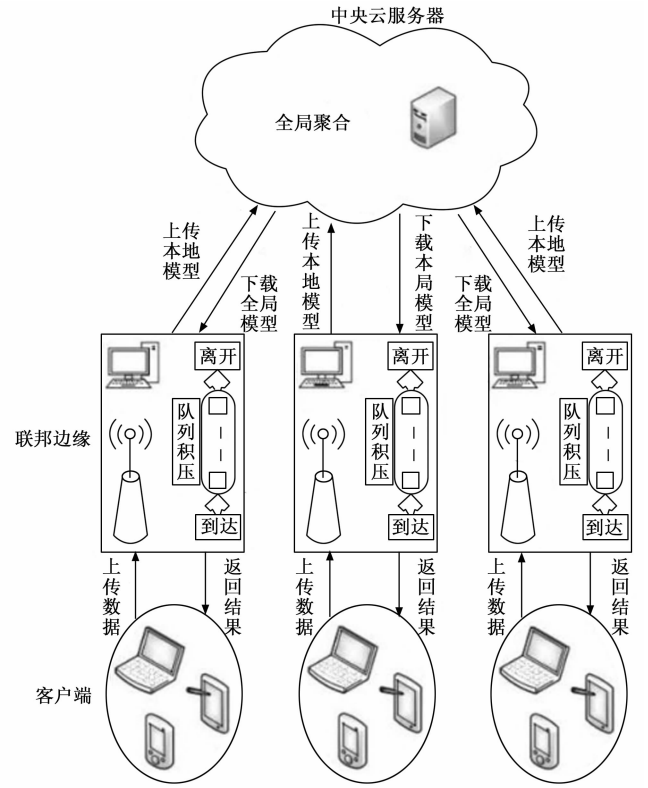

图 1 联邦边缘学习系统

2 Lyapunov 问题公式化

针对上一节提到的联邦边缘学习系统中联邦边缘的数据队列，进行进一步研究。考虑到在基于数据队列的系统中，确保队列的稳定性是必要的。随着参与训练的数据增加，模型精度也会提高。然而，由于队列积压不能无限增长，防止队列溢出是必要的。同时，队列的稳定性依赖于到达和离开过程。

因此，在这一部分，将基于 Lyapunov 优化理论[26]，根据队列的积压情况来确定最优选择客户端数，从而在模型精度和队列稳定性之间实现平衡。

数据队列在一个时隙的积压情况是由前一个时隙的到达和离开决定的。因此，数据队列可以表示为：

$$Q(t+1)=\max\{Q(t)+\lambda(t)-\mu(t),0\} \tag{1}$$

其中：$Q(t)$ 是 t 时隙队列积压量；$\lambda(t)$ 是 t 时隙到达队列数据量；$\mu(t)$ 是 t 时隙离开队列数据量。

对于每个联邦边缘，目标是在确保队列积压量不超过其最大长度的前提下，最大化模型精度。根据 Lyapunov 优化理论[27]，这个问题可以表示为：

$$P:\max:V\cdot U[n(t)]+Q(t)[\lambda(t)-\mu(t)] \tag{2}$$

$$\text{s.t.}\quad \lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}Q(t)\leqslant Q_{\max} \tag{3}$$

其中：$U[n(t)]$ 是 t 时隙选择客户端数为 $n(t)$ 时期望学习精度的效用函数；$Q(t)[\lambda(t)-\mu(t)]$ 是 Lyapunov 漂移；V 是用来权衡效用函数和 Lyapunov 漂移的平衡参数；T 是最大时隙数；$Q_{\max}$ 是队列最大长度。公式（3）是队列稳定性的约束。

由于每次客户端向联邦边缘传输的数据量相同，因此到达队列的数据量可以表示为：

$$\lambda[n(t)]=D\times n(t) \tag{4}$$

其中：D 是客户端传输的数据量。因此，P 的优化目标可以表示为：

$$n^*(t)\leftarrow\underset{n(t)\in K}{\operatorname{argmax}}\{V\cdot U[n(t)]+Q(t)\cdot\{\lambda[n(t)]-\mu(t)\}\} \tag{5}$$

其中：$n^*(t)$ 是 t 时隙最优选择客户端数；K 是客户端集。由于联邦边缘只有在通信条件足够好时才会将训练结果返回给客户端，且每个时隙的无线信道条件是随机的，因此 $\mu(t)$ 是一个随机值。

3 变分信息瓶颈问题公式化

针对上一节确定的最优选择客户端数，在这一部分，将使用信息瓶颈理论[28]进一步优化数据队列。信息瓶颈理论提供了一种高效的数据压缩方法，能够在保持数据队列大小不变的情况下，最大化数据的压缩和存储效率。

假设在每个客户端和联邦边缘都部署了编码神经网络和解码神经网络。每个客户端都携带数据集，数据集被划分为 B 个批次。记数据集、标签集、编码变量集、解码变量集分别为：

$$X=\{X_1,X_2,\cdots,X_B\} \tag{6}$$

$$Y=\{Y_1,Y_2,\cdots,Y_B\} \tag{7}$$

$$Z=\{Z_1,Z_2,\cdots,Z_B\}, \tag{8}$$

$$\hat{Y}=\{\hat{Y}_1,\hat{Y}_2,\cdots,\hat{Y}_B\} \tag{9}$$

其中：X_i、Y_i、Z_i 和 $\hat{Y}_i$ 分别是第 i 批次的数据集、

标签集、编码变量集和解码变量集（$i=1, 2, \cdots, B$）。

信息瓶颈公式[29]可以表示为：

$$R_{IB} = I(Z,Y) - \beta \cdot I(Z,X) \tag{10}$$

其中：$I(Z, Y)$ 是目标标签 Y 和解码向量 Z 之间的互信息；$I(Z, X)$ 是输入数据 X 和解码向量 Z 之间的互信息；β 是用来权衡 $I(Z, Y)$ 和 $I(Z, X)$ 的平衡参数。在信息瓶颈理论中，首要任务是提取最有用的部分信息，因此需要最大化 $I(Z, Y)$ 使解码向量更符合原始标签信息，并且最小化 $I(Z, X)$ 实现对信息更高的压缩。

$I(Z, Y)$ 可以表示为：

$$I(Z,Y) = \iint p(y,z)\log p(y \mid z)\mathrm{d}y\mathrm{d}z - \iint p(y,z)\log p(y)\mathrm{d}y\mathrm{d}z \tag{11}$$

由于公式（11）减号右边项对 z 积分为 1，对 y 积分变为 y 的熵，因此公式（11）可以简化为：

$$I(Z,Y) = \iint p(y,z)\log p(y \mid z)\mathrm{d}y\mathrm{d}z + H(Y) \tag{12}$$

$P(y, z)$ 的表达式可以根据编码器和马尔可夫链转移概率的性质 $Y \leftrightarrow X \leftrightarrow Z$ 导出，因此可以将公式（12）改写为：

$$I(Z,Y) = \iiint p(x,y)p(z \mid x)\log p(y \mid z)\mathrm{d}y\mathrm{d}z\mathrm{d}x + H(Y) \tag{13}$$

引入相对熵（Kullback-Leibler 散度）来解决 $P(y \mid z)$ 难以处理的问题，根据相对熵可以通过后验概率来间接推导互信息的上下界，从而满足信息瓶颈约束的条件。具体来说，利用相对熵的非负性这一重要特征，即：

$$KL[p(Y \mid Z), q(Y \mid Z)] \geqslant 0 \tag{14}$$

其中：$q(Y \mid Z)$ 是 $p(Y \mid Z)$ 的变分近似。根据定义可以得到：

$$\int p(y \mid z)\log p(y \mid z)\mathrm{d}y \geqslant \int p(y \mid z)\log q(y \mid z)\mathrm{d}y \tag{15}$$

进一步可以得到：

$$\iiint p(x,y)p(z \mid x)\log p(y \mid z)\mathrm{d}y\mathrm{d}z\mathrm{d}x \geqslant \iiint p(x,y)p(z \mid x)\log q(y \mid z)\mathrm{d}y\mathrm{d}z\mathrm{d}x \tag{16}$$

将公式（16）代入公式（13）可以得到 $I(Z, Y)$ 的下界：

$$I(Z,Y) = \iiint p(x,y)p(z \mid x)\log p(y \mid z)\mathrm{d}y\mathrm{d}z\mathrm{d}x + H(Y) \geqslant \iiint p(x,y)p(z \mid x)\log q(y \mid z)\mathrm{d}y\mathrm{d}z\mathrm{d}x + H(Y) \tag{17}$$

$I(Z, X)$ 可以表示为：

$$I(Z,X) = \iint p(x,z)\log p(z \mid x)\mathrm{d}x\mathrm{d}z - \int p(z)\log p(z)\mathrm{d}z \tag{18}$$

通常计算 Z 的边缘分布，即 $p(z)$，会很难，因此设 $r(z)$ 为这个边缘的变分近似。由于：

$$KL[p(Z), r(Z)] \geqslant 0 \tag{19}$$

根据定义可以得到：

$$\int p(z)\log p(z)\mathrm{d}z \geqslant \int p(z)\log r(z)\mathrm{d}z \tag{20}$$

因此，可以得到 $I(Z, X)$ 的上界：

$$I(Z,X) \leqslant \iint p(x,y)p(z \mid x)\log\frac{p(z \mid x)}{r(z)}\mathrm{d}x\mathrm{d}z \tag{21}$$

因此，信息瓶颈公式可以重新表示为：

$$R_{IB} = I(Z,Y) - \beta \cdot I(Z,X) \geqslant \iiint p(x,y)p(z \mid x)\log q(y \mid z)\mathrm{d}y\mathrm{d}z\mathrm{d}x - \beta \cdot \iint p(x,y)p(z \mid x)\log\frac{p(z \mid x)}{r(z)}\mathrm{d}x\mathrm{d}z \tag{22}$$

引入经验分布来近似 $p(x, y)$，即：

$$p(x,y) = \frac{1}{N}\sum_{n-1}^{N}\delta_{x_n}(x)\delta_{y_n}(y) \tag{23}$$

将经验分布和变分近似相结合，可以更准确地描述问题的不确定性，从而提高模型的泛化能力。其中，N 是每个批次解码变量的数量。因此，公式（22）可以改写为：

$$R_{IB} \geqslant \frac{1}{N}\sum_{n=1}^{N}\left[\begin{array}{l}\int p(z \mid x_n)\log q(y_n \mid z) - \\ \beta \cdot p(z \mid x_n)\log\dfrac{p(z \mid x_n)}{r(z)}\end{array}\right]\mathrm{d}z \tag{24}$$

接下来，客户端按照批次顺序将 X_i 输入到输出 Z_i 的编码神经网络中。z 的概率服从高斯分布，即：

$$p(z) = N(z \mid \boldsymbol{\mu}_i, \boldsymbol{\sigma}_i^2) \tag{25}$$

其中：N 是高斯分布，$\boldsymbol{\mu}_i$ 是编码函数的均值矩阵，$\boldsymbol{\sigma}_i^2$ 是编码函数的协方差矩阵。每个客户端都可以获得满足公式（25）的 $\boldsymbol{\mu}_i$ 和 $\boldsymbol{\sigma}_i^2$，并将它们传输到对应的联邦边缘。

每个联邦边缘基于接收到的 μ_i 和 σ_i^2 来生成重新参数化的变量集 $\hat{Z}$，即：

$$\hat{z} = \mu_i + \varepsilon\sigma_i^2 \quad \forall \hat{z} \in \hat{Z} \tag{26}$$

其中：$i \sim N(0, 1)$，并且 $\hat{z}$ 和 z 服从相同的分布。然后，每个联邦边缘将 $\hat{Z}$ 输入到输出 $\hat{Y}$ 的解码神经网络中。

由于 z 和 $\hat{z}$ 服从相同的分布，根据公式（22），计算 z 和 $\hat{z}$ 之间的互信息下界 $\hat{Y}_i$，记 $I(Z_i, \hat{Y}_i)_{\min}$，Z_i 和 X_i 之间的互信息上界，记 $I(Z_i, X_i)_{\max}$，然后根据互信息的上下界计算损失函数的下界，即：

$$\begin{cases} I(Z_i,\hat{Y}_i)_{\min} = \sum_{n=1}^{N}\sum_{m=1}^{M} p(z_{n,i} \mid x_{m,i}) \log_2 q(\hat{y}_{m,i} \mid z_{n,i}) \\ I(Z_i,X_i)_{\max} = \sum_{n=1}^{N}\sum_{m=1}^{M} p(z_{n,i} \mid x_{m,i}) \log_2 \dfrac{p(z_{n,i} \mid x_{m,i})}{r(Z_i)} \\ L_{min} = I(Z_i,\hat{Y}_i)_{\min} - \beta \cdot I(Z_i,X_i)_{\max} \\ \quad \forall z_{n,i} \in Z_i, \forall \hat{y}_i \in \hat{Y}_i \end{cases} \tag{27}$$

其中：M 和 N 分别是每个批次中（X_i，$\hat{Y}_i$）和 Z_i 的数量，$x_{m,i}$ 和 $\hat{y}_{m,i}$ 分别是 X_i 和 $\hat{Y}_i$ 的第 m 个数据，$z_{n,i}$ 是 Z_i 的第 n 个数据，p（$z_{n,i} \mid x_{m,i}$）是给定 $x_{m,i}$ 时 $z_{n,i}$ 的条件概率，q（$\hat{y}_{m,i} \mid z_{n,i}$）是对条件概率 p（$\hat{y}_{m,i} \mid z_{n,i}$）的变分近似，$r$（$Z_i$）是 Z_i 的边缘变分近似。根据信息瓶颈原理，通过最大化 I（Z_i，$\hat{Y}_i$）$_{\min}$ 可以使解码变量 $\hat{Y}_i$ 与标签信息 Y_i 更一致，通过最小化 I（Z_i，$_i$）$_{\max}$ 实现对信息更高的压缩。最后，每个客户端和联邦边缘都采用 Adam 算法来更新编码和解码神经网络的参数，并计算准确率：

$$acc = (1 - \frac{1}{B}\frac{1}{M}\sum_{i=1}^{B}\sum_{m=1}^{M} \hat{y}_{m,i} \oplus y_{m,i}) \times 100\% \tag{28}$$

4 仿真实验

在这一部分，将通过大量的仿真实验来比较关键性能指标，以验证数据队列优化方法的有效性。所有仿真实验均在华诚金锐 GS212S2 服务器上进行，服务器使用了全国产化的申威 3 231 处理器和 openEuler 系统。申威 3 231 处理器是我国自主研发的一款多核心处理器，基于 SW64 自主指令架构，集成 32 个 SW64 Core3B 核心、8 路 DDR4 存储控制器和 40lan PCIe4.0 接口，在处理器性能、访存性能以及 IO 性能方面表现突出。仿真实验基于 Python3.9，使用公开的 MNIST 数据集进行训练，假设系统中有 3 个联邦边缘，仿真结果是通过对 3 个联邦边缘的结果进行滑动平均得到的，仿真实验中使用的参数如表 1 所示。

表 1 仿真实验参数表

参数	值
客户端总数	50
数据队列最大长度	1 MB
Lypunov 平衡参数	10^9
总时隙数	400
客户端每次传输的数据量	7840 bits
变分信息瓶颈平衡参数	0.001
预期精度的学习率	1
预期精度的衰减率	−0.3

效用函数 U［n（t）］由[30] 中期望精度的学习曲线建模，如图 2 所示。当数据量为 α 时期望精度可以计算为：

$$E_{acc}(x) = 1 - l_{\text{rate}} \cdot \alpha^{d_{\text{rate}}} \tag{29}$$

其中：l_{rate} 是学习率，d_{rate} 是衰减率。

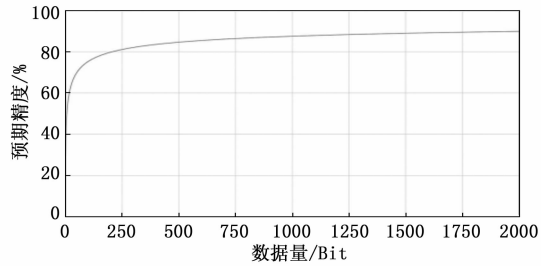

图 2 期望精度

为了评估提出的 Lyapunov 优化队列方法，将其与另外两种方法进行比较，即：

最大选择方法：联邦边缘在每个时隙接收所有客户端上传的数据[31-32]。

固定选择方法：联邦边缘在每个时隙都随机选择 5 个客户端上传数据[33]。

图 3 比较了不同方法的队列积压。黑色虚线表示队列的最大长度 $Q_{\max}$。可以看出，最大选择方法在每个时隙接收所有客户端上传的数据，因此队列积压迅速上升并超过 $Q_{\max}$，导致队列溢出而变得不稳定。采用固定选择方法时，每个时隙都随机选择 5 个客户端上传数据，队列积压远低于 $Q_{\max}$。虽然这样可以保证队列的稳定性，但用于训练的数据很少，且队列使用率很低。提出的 Lyapunov 优化队列方法的队列积压在前 100 个时隙都保持在 $Q_{\max}$，随后队列积压逐渐减少，直到所有客户端的电量耗尽。这是因为提出的方法在每个时隙选择最优数量的客户端上传数据，与其它两种方法相比，不仅确保了队列的稳定性，还能使更多的客户端参与训练。

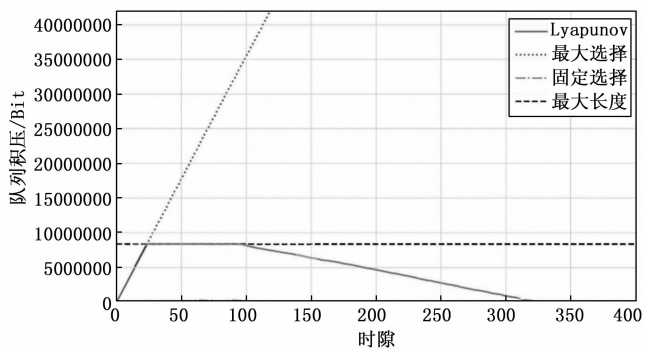

图 3 队列积压对比

因此，考虑将 Lyapunov 优化队列方法与提出的变分信息瓶颈方法相结合，进行后续的仿真实验。同时，还将引入一种对比方法，即：

Lyapunov＋随机选择方法：基于 Lyapunov 优化队列方法，在每个时隙随机选择最优数量的客户端上传数据[32-34]。

图 4 比较了不同方法参与训练的总数据量。提出的 Lyapunov＋变分信息瓶颈方法参与训练的总数据量为 47824000 Bits，而 Lyapunov＋随机选择方法和固定选

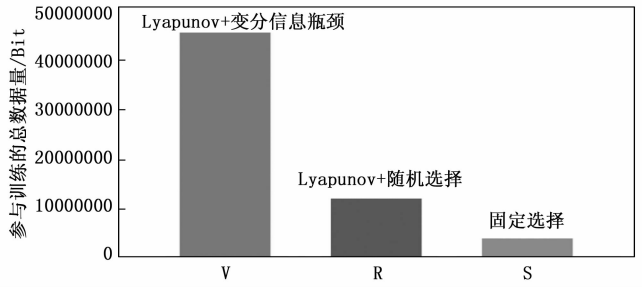

图 4　参与训练的总数据量对比

择方法参与训练的总数据量分别是 11956000 Bits 和 3820693 Bits。可以看出，提出的方法参与训练的总数据量远远大于其它两种方法。这是因为提出的方法不仅考虑了 Lyapunov 优化队列，还考虑对数据进行高效压缩，从而在相同队列容量下训练更多数据。此外，提出的方法在每个时隙还优先选择剩余电量较低的客户端上传数据，从而使更多数据得以参与训练。

图 5 比较了不同方法的模型精度。可以看出，提出的方法不仅在模型精度上优于其它两种方法，而且收敛速度也更快。这是因为在 Lyapunov 优化队列的基础上，提出的方法还考虑了高效数据压缩，并在每个时隙选择剩余电量较低的客户端上传数据，从而使更多数据参与训练，提高了模型精度并加快了收敛速度。此外，3 种方法在收敛性上基本相同，因为当训练数据量足够时，3 种方法都会实现收敛。

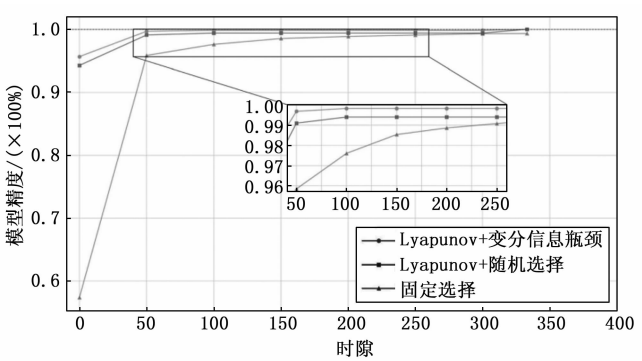

图 5　模型精度对比

图 6 比较了不同方法的损失值。与模型精度类似，由于提出的方法能够使更多数据参与训练，因此损失值低于其它两种方法，而且收敛速度更快。

上述结果表明，提出的方法要比其它几种方法的效果都要好。

另外，图 7 显示了变分信息瓶颈方法的准确度，在训练了很短的几个时隙之后准确度就能稳定达到 95% 以上。因此，可以看出变分信息瓶颈方法是一种高效的数据压缩方法。

5　结束语

在物联网大数据的背景下，传统的数据处理技术无

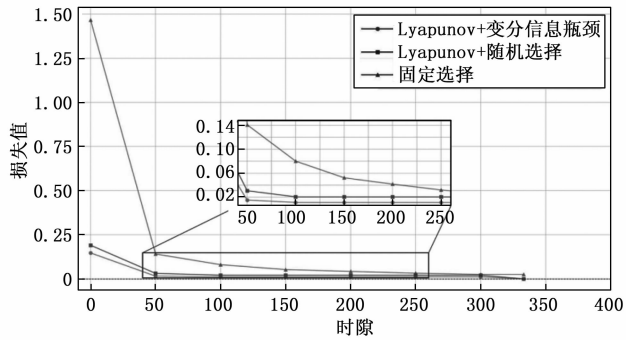

图 6　损失值对比

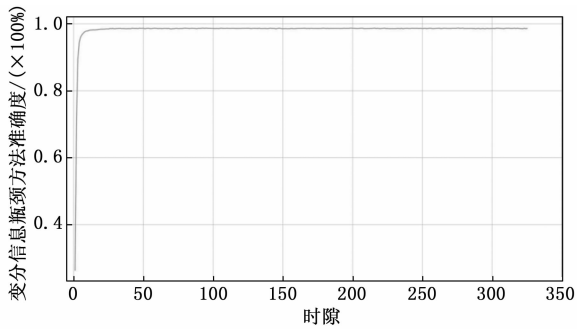

图 7　变分信息瓶颈方法准确度

法高效地处理和分析数据。针对上述问题，研究了物联网中的联邦边缘学习系统，并结合信息瓶颈理论提出了一种基于申威服务器的数据队列优化方法。该方法不仅缓解了客户端的计算负担，确保了联邦边缘数据队列的稳定性，提高了学习模型精确度，还提升了数据的存储和处理能力。通过仿真实验对该方法的性能进行了评估，结果表明，所提出的方法优于已知的基准方法。然而，本研究依然存在一些问题和不足，比如，虽然联邦边缘在训练结束后将训练结果发回给客户端，并删除已使用的训练数据，但是在客户端上传数据的过程中还是存在隐私泄露的风险。在后续的研究工作中，需要进一步考虑在联邦边缘学习系统中加强保护客户端隐私的机制，并使仿真实验更贴近实际应用场景。

参考文献：

[1] WANG S，TUOR T，SALONIDIS T，et al. When edge meets learning：adaptive control for resource-constrained distributed machine learning [C] // IEEE INFOCOM 2018-IEEE Conference on Computer Communications，2019，37：1205－1221.

[2] SATTLER F，WIEDEMANN S，MULLER K R，et al. Robust and communication-efficient federated learning from non-i. i. d. data [J]. IEEE Transactions on Neural Networks and Learning Systems，2020，31 (9)：3400－3413.

[3] 严　萍，张兴敢，柏业超，等. 基于物联网技术的智能家居系统 [J]. 南京大学学报（自然科学版），2012，48 (1)：26－32.

[4] 马士玲. 物联网技术在智慧城市建设中的应用 [J]. 物联网技术，2012，2 (2)：70－72.
[5] 刘小洋，伍民友. 车联网：物联网在城市交通网络中的应用 [J]. 计算机应用，2012，32 (4)：900－904.
[6] 杨海川. 基于物联网的智能家居安防系统设计与实 [D]. 上海：上海交通大学，2013.
[7] 方 晨，郭渊博，王一丰，等. 基于区块链和联邦学习的边缘计算隐私保护方法 [J]. 通信学报，2021，42 (11)：28－40.
[8] KONEN J，MCMAHAN B，RAMAGE D. Federated optimization：distributed optimization beyond the datacenter [J]. Mathematics，2015.
[9] LUO S，CHEN X，WU Q，et al. HFEL：joint edge association and resource allocation for cost-efficient hierarchical federated edge learning [J]. IEEE Transactions on Wireless Communications，2020，19 (10)：6535－6548.
[10] BONAWITZ K，EICHNER H，GRIESKAMP W，et al. Towards federated learning at scale：system design [J]. 2019.
[11] ZENG Q，DU Y，HUANG K. Wirelessly powered federated edge learning：optimal tradeoffs between convergence and power transfer [J]. IEEE Transactions on Wireless Communications，2022，21 (1)：680－695.
[12] JI Z，CHEN L，ZHAO N，et al. Computation offloading for edge-assisted federated learning [J]. IEEE Transactions on Vehicular Technology，2021，70 (9)：9330－9344.
[13] JEON J，PARK S，CHOI M，et al. Optimal user selection for high-performance and stabilized energy-efficient federated learning platforms [J]. Electronics，2020，9 (9)：1359.
[14] DU M，WANG K，XIA Z，et al. Differential privacy preserving of training model in wireless big data with edge computing [J]. IEEE Transactions on Big Data，2020，6 (2)：283－295.
[15] 孙一文. 基于信息瓶颈的强化学习泛化性问题研究 [D]. 天津：天津大学，2022.
[16] 马 儀，邵玉斌，杜庆治，等. 基于变分信息瓶颈多任务算法的多领域文本分类 [J]. 四川大学学报 (自然科学版)，2024，61 (3)：131－141.
[17] REN J，YU G，DING G. Accelerating DNN training in wireless federated edge learning systems [J]. IEEE Journal on Selected Areas in Communications，2021，39 (1)：219－232.
[18] SUN Y，ZHOU S，NIU Z，et al. Dynamic scheduling for over-the-air federated edge learning with energy constraints [J]. IEEE Journal on Selected Areas in Communications，2022，40 (1)：227－242.
[19] CAO X，ZHU G，XU J，et al. Optimized power control design for over-the-air federated edge learning [J]. IEEE Journal on Selected Areas in Communications，2022，40 (1)：342－358.
[20] FAN D，YUAN X，ZHANG Y，et al. Temporal-Structure-Assisted gradient aggregation for over-the-air federated edge learning [J]. IEEE Journal on Selected Areas in Communications，2021，39 (12)：3757－3771.
[21] MHAISEN N，ABDELLATIF A，MOHAMED A，et al. Optimal user-edge assignment in hierarchical federated learning based on statistical properties and network topology constraints [J]. IEEE Transactions on Network Science and Engineering，2022，9 (1)：55－66.
[22] ZOU Y，SHEN S，XIAO M，et al. Value of information：a comprehensive metric for client selection in federated edge learning [J]. IEEE Transactions on Computers，2024，73 (4)：1152－1164.
[23] 刘 峰，鄂海红. 基于海量数据的消息队列的性能对比与优化方案 [J]. 软件，2016，37 (10)：33－37.
[24] 李翠姣. 基于消息队列的分布式系统数据传输优化技术研究 [D]. 哈尔滨：哈尔滨工程大学，2015.
[25] 金海波，仲崇权. 基于排队论的实时以太网缓存队列优化算法 [J]. 大连理工大学学报，2012，52 (1)：95－99.
[26] NEELY M. Stochastic network optimization with application to communication and queuing systems [Z]. 2010.
[27] WU Q，WANG X，FAN Q，et al. High stable and accurate vehicle selection scheme based on federated edge learning in vehicular networks [J]. China Communications，2023，20 (3)：1－17.
[28] ALEMI A A，FISCHER I，DILLON J V，et al. Deep variational information bottleneck [J]. 2016.
[29] WU Q，KUAI L，FAN P，et al. Blockchain-Enabled variational information bottleneck for IoT networks [J]. IEEE Networking Letters，2024，6 (2)：92－96.
[30] FIGUEROA R L，ZENG-TREITLER Q，KANDULA S，et al. Predicting sample size required for classification performance [J]. Bmc Med In form Decis Mak，2012，12 (1)：8.
[31] KOTERA S，YIN B，YAMAMOTO K，et al. Lyapunov optimization-based latency-bounded allocation using deep deterministic policy gradient for 11ax spatial reuse [J]. IEEE Access，2021，9：162337－162347.
[32] BRACCIALE L，LORETI P. Lyapunov drift-plus-penalty optimization for queues with finite capacity [J]. IEEE Communications Letters，2020，24 (11)：2555－2558.
[33] MCMAHAN H B，MOORE E，RAMAGE D，et al. Communication-Efficient learning of deep networks from decentralized data [C] // Artificial Intelligence and Statistics (AISTATS)，Fort Lauderdale，Florida，USA，2017.
[34] PICANO B，FANTACCI R，HAN Z. Aging and delay analysis based on lyapunov optimization and martingale theory [J]. IEEE Transactions on Vehicular Technology，2021，70 (8)：8216－8226.