智能物联网应用的个性化联邦学习：基于云边缘的框架

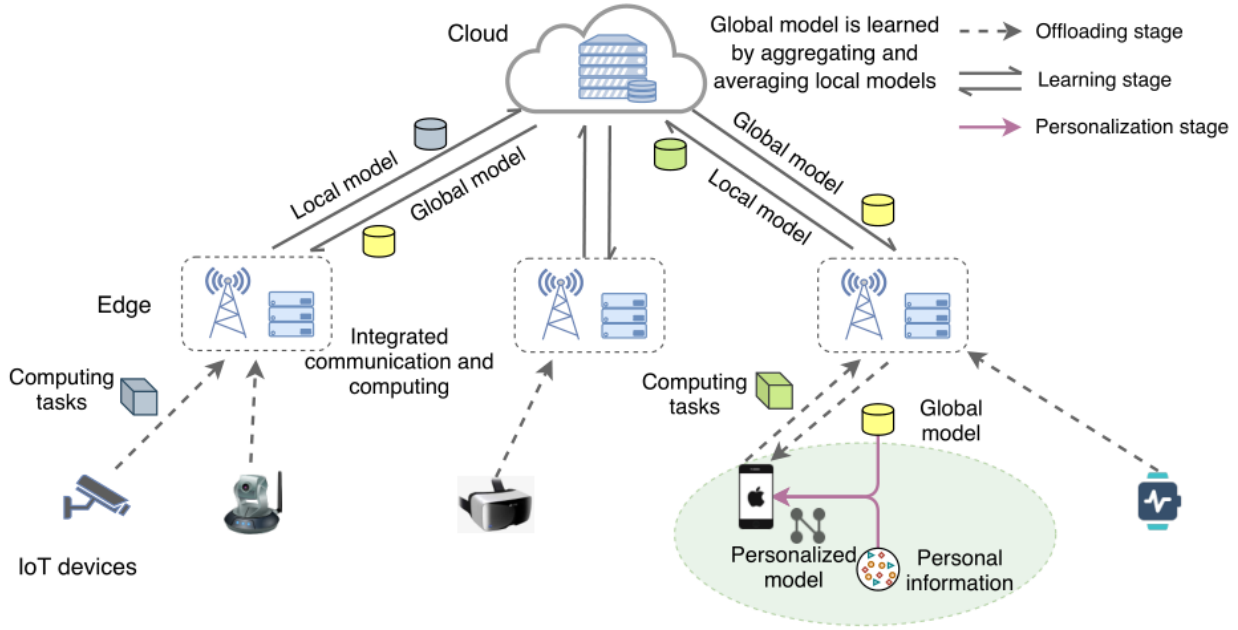
在本文中，提出了一个名为PerFit的协同云边缘框架，用于个性化联合学习，以整体方式减轻物联网应用中固有的设备异构性、统计异构性和模型异构性。
求助于边缘计算，每个物联网设备可以选择将其计算密集型学习任务卸载到边缘，从而满足快速处理能力和低延迟的要求。此外，边缘计算可以通过在本地附近存储数据(例如，在家中智能家居应用的智能边缘网关中)来缓解隐私问题。此外，还可以采用差分隐私和同态加密等隐私和安全保护技术来提高隐私保护级别。对于统计和模型异构性，该框架还使得终端设备和边缘服务器能够在云-边缘范例中在中央云服务器的协调下联合训练全局模型。在通过联合学习训练了全局模型之后，在设备侧，然后可以采用不同种类的个性化联合学习方法来实现针对不同设备的个性化模型部署，以适应它们的应用需求。

在物联网方面联邦学习面临的挑战：

1.设备的异构性，例如存储，计算和通信能力和通信安全；
2数据异质性，非IID（非独立同分布）性质；
3.模型异质性，即不同设备想要根据其应用环境定制其模型的情况；

作者提出以整体方式应对异构挑战，提出的PerFit框架采用云边缘架构，可在IoT设备附近带来必要的按需计算能力，这样每个设备可以选择将其密集的计算任务卸载到边缘（家

庭中的边缘网关，办公室中的边缘服务器或户外的5G 移动边缘计算（MEC）服务器）来满足高处理效率和低时延的要求。
然后在终端设备、边缘服务器和远程云之间使用联邦学习FL，这可以通过汇总来自IoT用户的本地计算模型来联合训练共享的全局模型，同时将所有敏感数据保留在设备上。接着采用个性化的FL方法来微调每个设备的学习模型，以解决异质性问题。



PerFit中的协作学习过程主要包括以下三个阶段：

1. 卸载阶段： 当边缘是可信赖的时候（如家中的边缘网关），IoT设备可以将其整个学习模型和数据样本卸载到边缘以进行快速计算；当边缘不可信赖时，设备用户进行模型划分，将输入层和数据样本保存在本地设备上，将其余的层卸载到边缘来进行设备-边缘到协同计算。
2. 学习阶段： 设备和边缘基于个人数据样本共同计算本地模型然后将其上传到云服务器；云服务器收集、平均各个边缘上传的本地模型信息得到一个全局模型，并将其回传给边缘。重复这个信息交换过程直至收敛，便可得到一个高质量的全局模型并将其传给边缘用来做个性化。
3. 个性化阶段：每个设备会基于全局模型信息和自己的个性化信息训练个性化模型，此阶段的具体学习操作取决于所采用的个性化联邦学习机制。

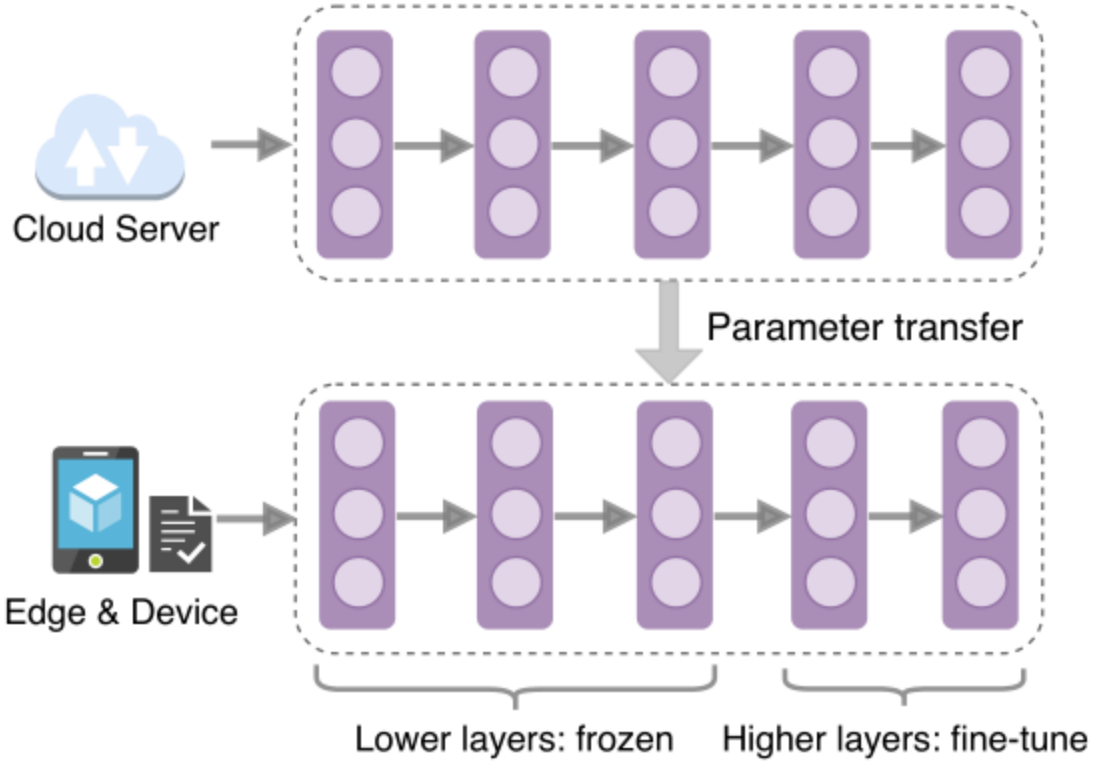
个性化联邦学习机制

1.联邦迁移学习（Federated Transfer Learning）：

FL的目标是将训练过的参数从源域转移到目标域，FTL的主要思想是将全局共享模型分发到IoT设备上用以个性化来减轻FL中的统计异质性。

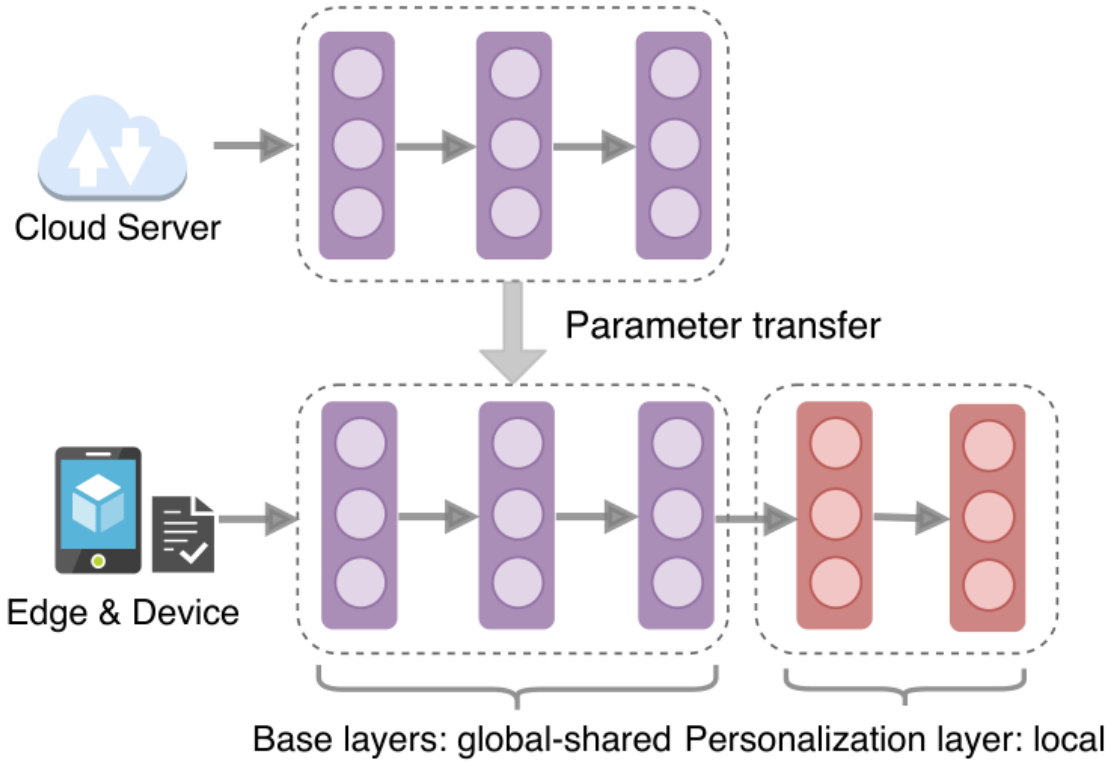
FTL中主要的个性化方法：

(1)首先通过传统的联合学习来训练全局模型，然后将全局训练的模型传送回每个设备。因此，每个设备能够通过用其本地数据提炼全局模型来构建个性化模型。为了减少训练开销，只微调指定层的模型参数，而不是重新训练整个模型。全局模型的较低层中的模型参数可以被转移并直接重新用于局部模型，因为深层网络的较低层关注于学习公共和低级特征。而较高层中的模型参数应该随着它们学习到为当前设备定制的更多具体特征而用本地数据进行微调。



(2)采用不同的方式通过联合迁移学习来执行个性化。将深度学习模型视为基础+个性化层，基础层充当共享层，使用现有的联合学习方法(即，FedAvg方法)以协作方式对其进行训练。而个性化层在本地被训练，从而能够捕获物联网设备的个人信息。以这种方式，在联合训练过程之后，全球共享的基础层可以被转移到参与的物联网设备以进行构建

他们自己的个性化深度学习模型及其独特的个性化层。因此，FedPer能够捕获特定设备上的细粒度信息，以进行更好的个性化推断或分类，并在一定程度上解决统计异构性。



2.联邦元学习（Federated Meta Learning）

在元学习中，模型由元学习者训练，该元学习者能够学习大量相似的任务，并且训练后的模型的目标是从少量新数据中快速适应新的相似任务。元学习可以与任何基于梯度训练的模型表示灵活地结合，还可以通过少量的数据样本快速学习和适应。

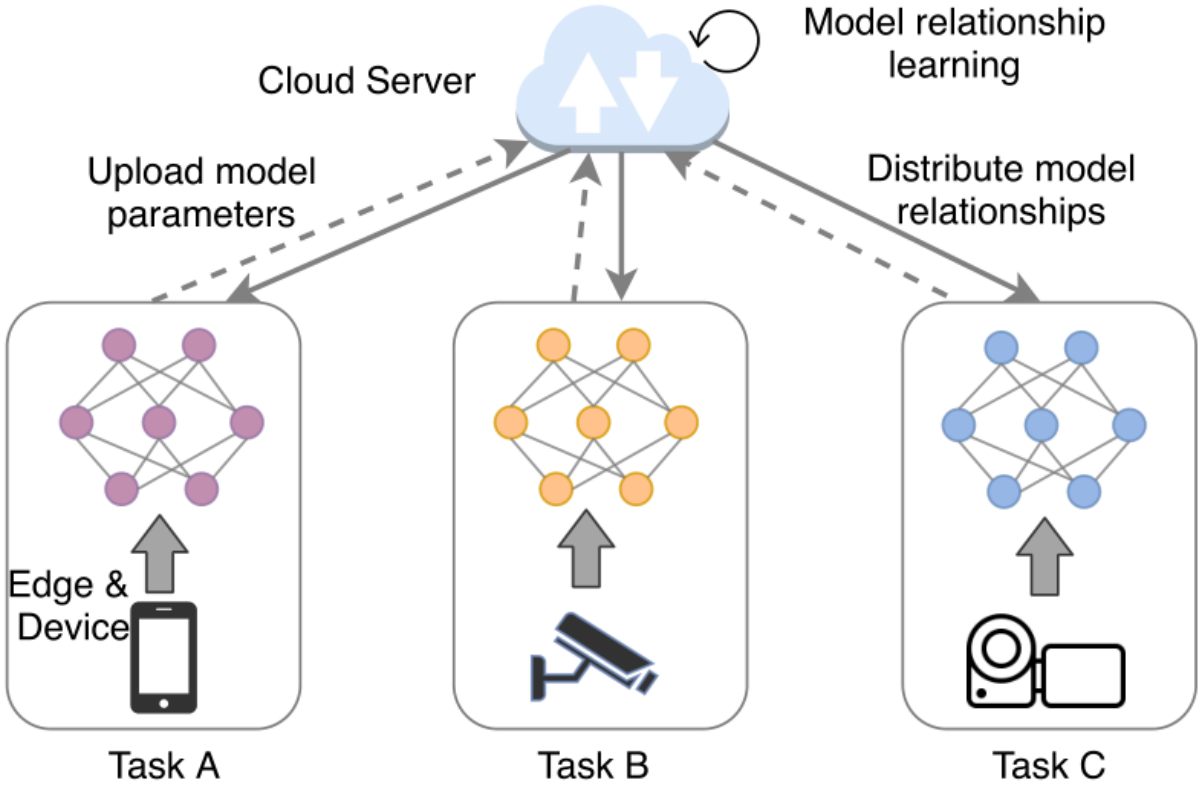
由于联邦元学习方法经常利用复杂的训练算法，因此其实现复杂度高于联合转移学习方法。但是，通过联邦元学习获得的学习模型更加健壮，并且对于那些数据样本很少的设备可能非常有用。

3.联邦多任务学习（Federated Multi-task Learning）：

FTL和FML旨在通过微调的个性化方法在IoT设备之间学习相同或相似任务的共享模型；而FMTL旨在同时学习针对不同设备的不同任务，并在没有隐私风险的条件下寻找模型之间的联系。通过这些联系，每个设备可以得到其他设备的信息，且为每个设备学习的模型始终是个性化的。

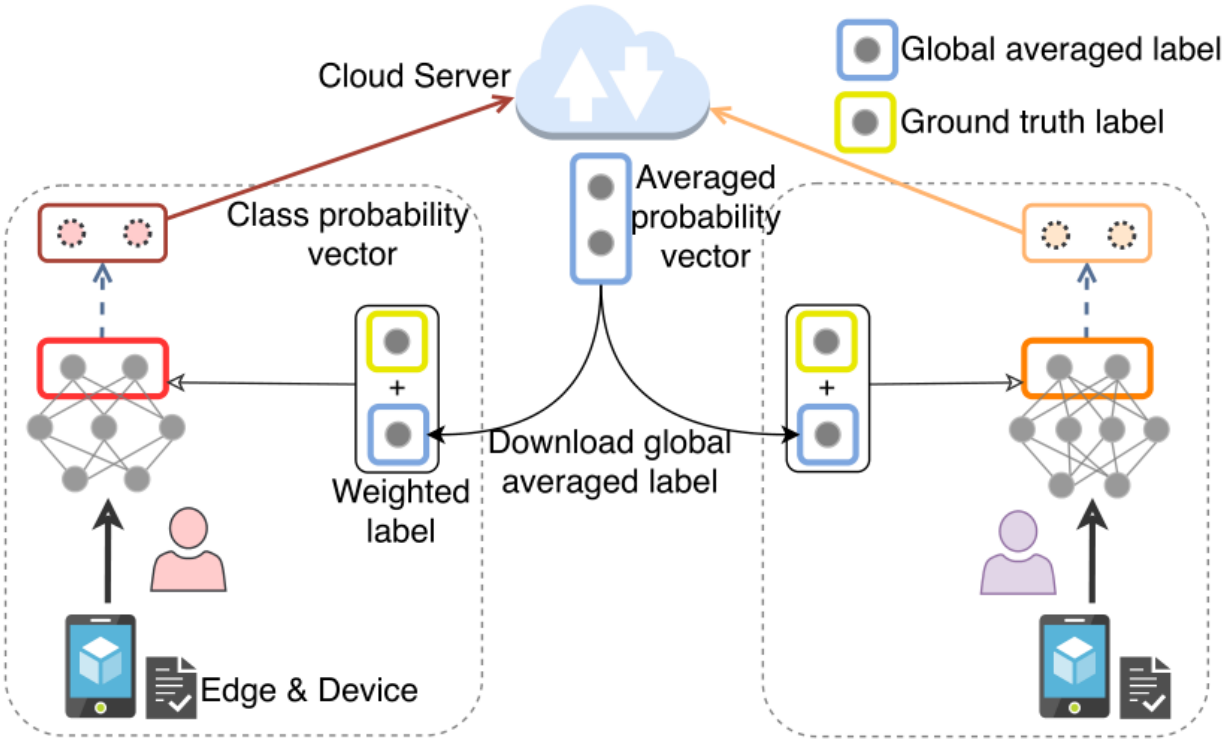
在FMTL训练中，云服务器基于物联网设备上传的模型参数来学习多个学习任务之间的模型关系；然后，每个设备都可以使用其本地数据和当前模型关系更新其自己的模型参

数；通过对云服务器中的模型关系和每个任务的模型参数进行交替优化，FMTL使参与的物联网设备协作培训其本地模型，从而减轻统计异质性并得到高质量的个性化模型。



4.联邦蒸馏（Federated Distillation）：

利用知识蒸馏的能力使每个参与者自主设计自己模型的新型FL模型。每个客户都需要将其学到的知识转换为标准格式，其他人可以在无需共享数据和模型架构的情况下理解这些格式；然后中央服务器收集这些知识以计算共识，并将共识进一步分发给参与的客户端。知识转换步骤可以通过知识蒸馏来实现。通过这种方式，云服务器对每个数据样本的类概率进行汇总和平均化然后分发给客户以指导他们的更新。



5.数据增强(Data Augmentation）：

由于用户个人生成的数据自然呈现出高度偏斜且非IID的分布，这可能会大大降低模型的性能，因此出现了致力于数据增强以促进个性化联邦学习的新兴作品。一种新型数据共享策略，即分配少量的包含从云到边缘客户端的类之间的均匀分布的全局数据，但是，直接将全局数据分发给边缘客户端会带来很大的隐私泄漏风险，需要使用此方法在数据隐私保护和性能改进之间进行权衡。而且，全局共享数据和用户本地数据两者之间的分布差异也会带来性能下降。

为了纠正不平衡且非IID的本地数据集而又不损害用户隐私，采用了一些过采样技术和具有生成能力的深度学习方法。Jeong提出了联邦增强（FAug）：每个边缘客户端识别其数据样本中缺少的标签（目标标签）；然后将这些目标标签的少量种子数据样本上载到服务器；服务器对上传的种子数据样本进行过采样，然后训练一个生成对抗网络（GAN）；最后，每台设备都可以下载经过训练的GAN生成器，以补充其目标标签，直到达到平衡的数据集为止。

通过数据增强每个客户端都可以基于生成的平衡数据集训练更加个性化和准确的分类或推断模型。值得注意的是，FAug中的服务器应该是可信任的，以便用户愿意上传其个人数据。

结论：

本文提出了一个云边缘架构中的个性化联邦学习框架 PerFit，用于具有数据隐私保护的智能物联网应用。PerFit 能够通过聚合来自分布式物联网设备的本地更新并利用边缘计算的优点来学习全局共享模型。

为了解决物联网环境中的设备、统计和模型的异构性，PerFit 可以自然地集成各种个性化联邦学习方法，从而实现物联网应用中设备的个性化处理并增强性能。