

请发明人员自行

复印和保留复本

**发**

**明**

**报**

**告**

**书**

**(1)**

|  |  |  |  |
| --- | --- | --- | --- |
| 提交日： 年 月 日 | 将以本报告书为基础进行专利检索和申请，以下**\***部分为必填项目。 | | |
| **\***发明名称(**中+英**)： 一种分割学习（模型拆分）辅助的联邦学习网络  A Split Learning (Model Splitting) Aided Federated Learning (SL-aided FL) Network | | | |
| **\***发明概要：(请用100字左右总结该发明)  此方案提出了一种新型联邦学习网络结构，将 UE周边设备作为辅助设备（Assistant Equipment，AE）， UE和AE间通过sidelink与连接进行split learning。完整的模型被拆分，部分模型的训练由AE完成训练。充分利用了AE的计算或通信能力。此外，该方法结合了split learning和federated learning，数据隐私保护也进一步提升。 | | | |
| **\***Inventor Information Sheet(超过两位请另填Excel附件)  1. 发明人是指对本发明区别于现有技术的部分作出实质性贡献的人员。  2. 发明人员中包括公司外人员，或者本发明是与其他单位联合研发时，请明确注明。   |  |  |  | | --- | --- | --- | |  | 第一发明人 | 第二发明人 | | 英文名（拼音） | Ce ZHENG |  | | 中文名 | 郑策 |  | | 员工号 | 700027697 |  | | 所属单位/部门/联系电话 |  |  | | 居住地 (做出发明时所在的居住地) | Beijing, China |  | | 国籍 | China |  | | 中文通信地址 | 北京市朝阳区新源南路1号平安国际金融中心商业栋3层301室 100027 |  | | 英文通信地址 | Room 301, 3/F, Commercial Podium, Ping An International Financial Center，No.1 South Xinyuan Rd. Chaoyang District, Beijing 100027 Beijing China |  | | 电子邮箱 |  |  | | 身份证号（中国籍的第一发明人须填写） | 231121199110150635 | N/A | | | | |
| **\***技术领域(e.g. 移动通信LTE-A)：  **\***涉及具体项目：□是，项目名称 □否  **\***涉及技术标准：□是，标准名称 □否  发明独特性：□唯一解法 □少数解法中的一种 □多种解法中的一种 | | | **\***研发状态：  □创意阶段  □研发/试作阶段  □产品化阶段，机型 | |
| **\***公开计划：  □无 □有，预定日期  公开方式：  □展示 □销售 □订货 □出厂 □发表论文 □其它 | | **\***现有技术调查：□未调查 □已调查  检索结果：（与发明重点相关的参考文献或专利申请号。空间不够时请添加附件） | | |

以下评价表格为**部门填写**：

|  |  |  |  |
| --- | --- | --- | --- |
| 统括课长 | **\***技术及市场角度综合评价（请领导填写） | 统括部长 | 知产负责人 |
|  | □重点，原因  □一般  □不申请，原因  （e.g.无保护必要，可对外公开；以商业秘密等其他形式保护） |  |  |
| 意见： | | | |

**发**

**明**

**报**

**告**

**书**

(2)

**发明的重点**

专利说明书中的[权利要求范围]（CLAIM）的栏目内容。发明的要点（CLAIM）**至少7项**，其中，第1项发明重点应当包括对最主要创新点的概括，其他项发明重点可以是在第1项基础上附加的创新点的概括，并且请对每一项发明重点中包含的创新点所带来的技术问题及效果进行说明。

|  |  |
| --- | --- |
| **权利要求** | **技术问题及效果** |
| 1. **一种 分割学习（模型拆分）辅助的联邦学习网络**：   该网络结构由三部分组成：gNB基站，Standalone Member和Non-Standalone Member。其中，  Standalone Member 为独立的用户设备（UE）；  Non-Standalone Member 由UE用户设备和AE辅助设备共同组成，二者通过sidelink连接。 | 此可能为本专利的主要创新点（联邦学习和分割学习的网络结构已经存在，**但是没有人考虑过用sidelink将两个设备连接起来。两个设备之间进行split learning。然后UE与AE作为一个整体（Member或Client）参与FL。** |
| 1. **一种配置在UE端的设备**，主要用于：   对与自身状态相关的信息 进行估计并上报至gNB；  作为Standalone Member, 接收来自于gNB的全局模型；  基于本地数据，进行本地的模型训练与模型更新；  上传更新的本地模型至gNB；  或者  与gNB的指定的AE共同组成Non-Standalone Member，接收来自于gNB的全局模型及split point或对应的UE model；  基于本地数据，与AE共同进行本地的模型训练，即分割学习。其中，UE执行运行UE model部分（正向传播算法），并将所得的中间数据#1传输至AE；接收AE传回的中间数据#2，更新根本模型（反向传播算法）；    将更新后的UE model按照gNB的指示，发送至gNB或AE。 |  |
| 1. **一种配置在AE端的设备**，主要用于：   对与自身状态相关的信息 进行估计并上报至gNB；  与gNB的指定的UE共同组成Non-Standalone Member，接收来自于gNB的全局模型及split point或对应的AE model；  与UE进行分割学习。接收UE传来的中间数据#1，作为AE model的输入，进行AE model的训练（正向和反向传播算法），并将得到的中间数据#2传回UE；  将更新后的AE model按照gNB的指示，发送至gNB或UE。 |  |
| 1. **一种配置在gNB端的设备**，主要用于：   接收来自于UE和AE上报的状态信息（和），选择参与联邦学习的Standalone Member (UE)和Non-Standalone Member (UE和AE)；决定Non-Standalone Member的Split point；分配uplink, sidelink, 以及可能会分配downlink的资源（如果对AE和UE分别发送对应的模型的话）；  广播全局模型并告知Non-Standalone Member中的UE和AE其Split Point 或者 对Non-Standalone Member中的AE和UE分别发送对应的模型；  告知AE和UE模型上传流程：各自上传模型还是由其中一方完成拼接后上传；  接收来自于Standalone Member (UE)的本地模型；  以及  接收来自于UE 和/或者 AE的本地模型(UE model, AE model或者完整的model)，进行模型的拼接（如果接收不是完整的model）；  全局模型的聚合； |  |
| 1. 根据权利要求2，UE端的设备进行自身状态相关的信息 -- 的估计，其中，包括但不局限于   UE与gNB之间信道的状态信息（如SINR, RSRP等）  UE自身的计算和存储能力信息（如CPU占用率，内存等）  本地数据信息（如参与模型训练的样本数量，样本维度等）  电量，位置信息，移动性(如速度，方向，在本地停留时间)等  UE与周边设备的信道（sidelink）状态信息  并将进行上报至gNB； |  |
| 1. 根据权利要求2，AE端的设备进行自身状态相关的信息 -- 的估计，其中，包括但不局限于   AE与gNB之间信道的状态信息（如SINR, RSRP等）  AE自身的计算和存储能力信息（如CPU占用率，内存等）  电量，位置信息，移动性(如速度，方向，在本地停留时间)等  AE与周边UE的信道（sidelink）状态信息  并将进行上报至gNB； | XXXXXXXXX |
| 1. 根据权利要求2，3   Non-Standalone Member中的UE与AE进行Split Learning。完整的model根据split point分割为UE model 和 AE model。UE与AE通过sidelink相连接。其流程如下：  1). UE首先执行运行UE model部分（正向传播算法），并将所得的中间数据#1传输至AE；  2). AE将接收到的中间数据#1 作为AE model的输入，首先运行正向传播算法得到结果。其次根据结果，运行反向传播算法更新AE model，并得到中间数据#2。将中间数据#2传回UE；  3). UE将根据中间数据#2运行反向传播算法，更新UE model。  其中每轮全局训练 可能会 重复1). – 3) K次，K的取值不在本专利的讨论范围内。 |  |
| 1. 根据权利要求4，gNB端的设备接收来自于UE和AE上报的状态信息（和），选择参与联邦学习的Standalone Member (UE)和Non-Standalone Member (UE和AE)；决定Non-Standalone Member的Split point。可细分为以下几种场景：   1). UE#i已确定作为Non-Standalone Member的UE参与联邦学习 且 UE#i和AE之间的Split point也已经确定，但对应于UE#i的AE还未确定  2). UE#i已确定作为Non-Standalone Member的UE参与联邦学习 且 对应的AE#j也已确定，但UE#i和AE#j之间Split point还未确定  3). UE#i已确定参与联邦学习，但未确定是作为Standalone Member 还是和其他UE组成Non-Standalone Member（即是否需要进行Split learning）还未确定  4). UE#i未确定是否参与联邦学习 |  |
| 1. 针对权利要求8中的   1) UE#i已确定作为Non-Standalone Member的UE参与联邦学习 且 UE#i和AE之间的Split point也已经确定，但对应于UE#i的AE未决定。  此时，对于AE的选择需要考虑UE#i与各可选AE之间的**sidelink**。  如  a). 可选AE对应的**sidelink**的RSRP，RSRQ或者SNR等参数中的一个或多个需大于一定阈值；  b). 可选AE对应的sidelink需满足：对于给定Split point对应的 **中间数据#1**（参考权利要求2和3），在给定的时间范围（）内完成从 UE到AE 的成功传输  和/或 **中间数据#2**在给定的时间范围（）内完成AE到UE之间的成功传输  等等 |  |
| 1. 针对权利要求8中的   2). UE#i已确定作为Non-Standalone Member的UE参与联邦学习 且 对应的AE#j也已确定，但UE#i和AE#j之间Split point不确定  此时，对于Split point的选择，需要考虑UE#i和AE#j间的**sidelink**状态。  如  a). 在该**sidelink**的状态下，所选Split point应满足：该Split point对应的 **中间数据#1** 在给定的时间范围（）内完成从 UE到AE的成功传输 和/或 **中间数据#2**在给定的时间范围（）内完成AE到UE之间的成功传输； |  |
| 1. 针对权利要求8中的   3). UE#i已确定参与联邦学习，但未确定是作为Standalone Member 还是和其他AE组成Non-Standalone Member（即是否需要进行Split learning，选取哪个AE，哪个Split point）还未确定  此时，需要考虑UE#i和其相连接的可选AE的**sidelink**状态。  如  a). 如果 AE#j对应的**sidelink**的RSRP，RSRQ或者SNR等参数中的一个或多个大于给定阈值的情况下，可以选取该AE#j与UE#i作为Non-Standalone Member参与FL。反之，如果不存在满足该条件的AE，可以选取UE#i作为Standalone Member参与FL；  b). 如果UE#i与AE#j 通过**sidelink**连接 组成Non-Standalone Member。在该**sidelink**状态下，存在一个或多个Split point，使该Split point对应的 **中间数据#1** 在给定的时间范围（）内完成从 UE到AE的成功传输 和/或 **中间数据#2**在给定的时间范围（）内完成AE到UE之间的成功传输，可以选取该AE#j与UE#i作为Non-Standalone Member，在该Split point处进行模型分割执行SL。反之，如果不存在满足该条件的AE，可以选取UE#i作为Standalone Member参与FL；  等 |  |
| 1. 针对权利要求8中的   4). UE#i未确定是否参与联邦学习  此时，需要考虑UE#i和其相连接的可选AE的**sidelink**状态.  如  a). 如果 AE#j对应的**sidelink**的RSRP，RSRQ或者SNR等参数中的一个或多个大于给定阈值的情况下，可以选取该UE#i与该AE#j作为Non-Standalone Member参与FL。  b). 如果UE#i与AE#j 通过**sidelink**连接 组成Non-Standalone Member。在该**sidelink**状态下，存在一个或多个Split point，使该Split point对应的 **中间数据#1** 在给定的时间范围（）内完成从 UE到AE的成功传输 和/或 **中间数据#2**在给定的时间范围（）内完成AE到UE之间的成功传输，可以选取该AE#j与UE#i作为Non-Standalone Member，在该Split point处进行模型分割执行SL。 |  |
| 1. 针对权利要求8-12，由于UE#i与AE#j中的sidelink随时间变化而变化。因此，针对权利要求中的1)-4)，即UE#i是否参与联邦学习，作为Standalone Member还是Non-standalone member，与哪个AE组成Non-standalone member，Split point的位置，需要根据sidelink的变化而进行动态的决策。 |  |

（超出10项请追加）

**中英文关键词：**

FL --- Federated Learning 联邦学习

SL --- Split Learning 拆分学习/分割学习（CCSA的白皮书输入报告中，使用的是分割学习）

UE --- User Equipment 用户设备

AE --- Assistant Equipment 辅助设备

专利说明书中描述[现有技术]的栏目。请尽可能引用专利公报、学术论文、标准提案等来说明现有技术，例如**概括若干文献中的技术方案**并说明其欠缺点。

**现有技术及其问题点**

|  |  |  |  |
| --- | --- | --- | --- |
| **General background** | 1. **Federated Learning** 2. **V2X** | | |
| **具体内容：** | | | |
| **TS/TR SPEC**  **（if possible）** | **SPEC No.** | | **TR 38.331 TS 22.186** |
| **Chapter No.** | | **6.3.2** |
| **具体内容（直接粘贴-与你技术方案相关的）-<如果标准化，会对标准的哪个章节有影响？>：**   1. **TR 22.874** AMMT   7. Distributed/Federated Learning over 5G system | | | |
| **Meeting/agreements** | **Meeting No.** | |  |
| **Meeting date** | |  |
| **SPEC No.** | |  |
| **具体内容（直接粘贴）：** | | | |
| **Techical Problem** | | | |
|  | | | |
| **Existing solution for above problem** | | **Other companies’ proposals (one or more)?** | |
| **Company/Tdoc No.** |  |
|  | |  | |
|  | | | |
| **Disadvantage or problem of above existing solution** | | | |
| **解决方案具体说明(可把要点粘贴，并作必要的解释说明)：** | | | |
| **Summary of potential standardized points(你认为可能标准化的点列出来，以标准的形式，like proposal)**  SA1 和SA在Release-19提的WID “Study on AI/ML Model Transfer Phase 2” 。文件是SP-220439 “**Distributed AI training/inference based on direct device connection**”. The objectives are to study new use cases and potential service and performance requirements to support efficient AI/ML operations using direct device connection , including:Distributed AI training/inference based on direct device connection, e.g. traffic KPIs, different QoS and functional requirements on slidelink transmission.Charging and security aspects. **主要是网络结构的创新,利用了sidelike** | | | |

**发**

**明**

**报**

**告**

**书**

(3)

1. 请写明为了实现此发明、你认为最佳的实施状态是怎样的（包括发明所要使用的设备、系统及重要部件等全部内容）。
2. 请尽可能详细写明此发明重要部件的结构、运行方式、作用等。
3. **请写出由此发明引申出的、其他相近似的方案的至少两个示例。**
4. **对于所列出的公式请说明公式整体含义及各个参数的具体含义，并给出可选的其他公式示例。**

注 参照图纸、图表、流程图等资料时，请编号进行说明、描述；如果有技术报告等材料，请充分利用，加以辅助说明。

发明的具体说明

专利说明书中关于[实施实例]的部分。请按照下列顺序加以详细说明。

Beijing

1. **发明背景**
   1. **联邦学习基本原理与网络结构**

联邦学习网络结构如图1所示

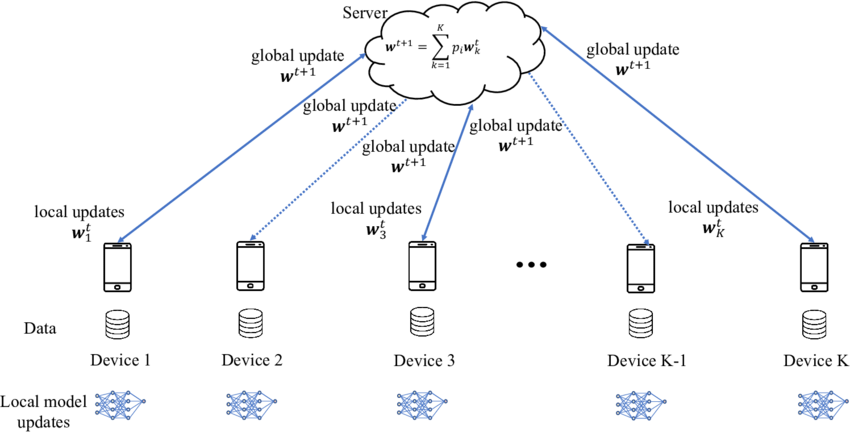


图1. 联邦学习网络结构

各Device或UE通过无线信道接入Server或gNB进行通信，将本地所学习到的模型上传至Server，Server进行聚合后再下发至各设备。具体流程如下：

1. 各设备接入Server，通过Server下行链路传输，获取初始全局学习模型()：
2. 各设备使用存储在本地的数据进行学习，完成一次本地模型更新的迭代:

,

其中表示梯度，表示设备的损失函数

1. 各设备通过上行链路，将学习到的本地模型（）或梯度（）上传至Server；
2. Server将收集到的来自各设备的本地模型进行聚合，完成全局模型的更新；

,

其中为来自于各设备的本地模型的权值, 通常设置为 ；

1. Server将更新后的全局模型再次下发到各设备，运行 1）--4） 直至模型收敛。

然而，该网络结构有其相应的局限性：

1. 设备通常通过与gNB相连上传至Server，而gNB的覆盖范围有限，对于覆盖范围外的用户或者设备无法提供服务
2. gNB覆盖范围内存在一些通信速率较低的地区，对于该地区的用户，即使在覆盖范围内，服务质量无法保证
3. gNB上下行链路通信资源有限，如频率资源，载波个数等，无法支持过多用户的同时接入。
   1. **分割学习基本原理与网络结构**

分割学习模型示如图2所示

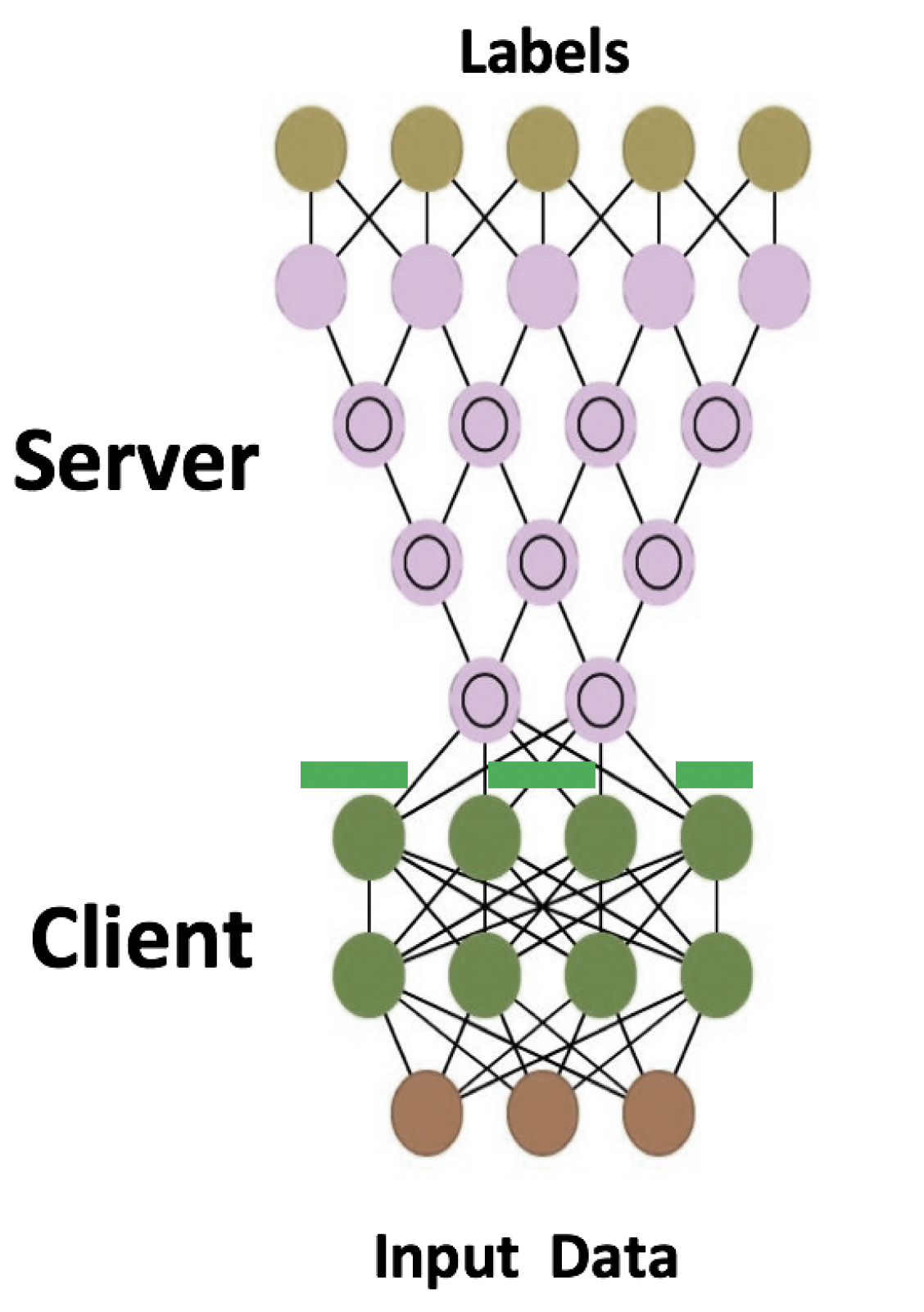


图2. 分割学习模型示意图0

神经网络模型被分割为两部分，分别为位于Client的前（几）层模型和位于Server的后（几）层模型。我们分别称之为Client model和Server model。

其训练过程可分为两步0：步骤 a). 确立模型拓扑结构; 步骤 b). 模型训练。

a). 如图2所示，神经网络模型被分割为两部分，分别为位于UE的前（几）层模型和位于Server的后（几）层模型。我们分别称之为Client model和Server model。 训练开始时，两部分模型均随机初始化训练参数。

b). 在训练过程中，Client基于本地数据对本地模型进行前向计算，并将本地标签（Label） 和 前向计算得到的中间数据（即分割层的输出张量）和发送给Server。Server在得到中间数据后继续进行前向计算，并根据上传的标签进行反向梯度计算，得到Server model对应的梯度。Server将分割层的梯度反向传回Client，Client继续进行反向计算得到相应的梯度。Client和Server根据各自得到的梯度进行模型参数更新。如此循环往复，直到模型收敛。

**注1：Split Learning 可以在单个client和server之间进行，也可以在多个client和server之间进行**。

图3给出了SL在多用户场景下的示意图所示：

1. Server首先跟Client#1建立连接，Client#1基于本地数据与Server进行一次或多次SL，完成Client model和Server model的模型参数更新；
2. 随后Server再跟Client#2建立连接。同时，Client#2需要获取Client#1训练得到的模型参数进行Client Model的本地更新[注2];
3. Client#2基于本地数据与Server进行一次或多次SL，完成Client model和Server model的模型参数更新；
4. 类似1）-3），进行Client#2 --> Client#3-->… 至模型收敛。

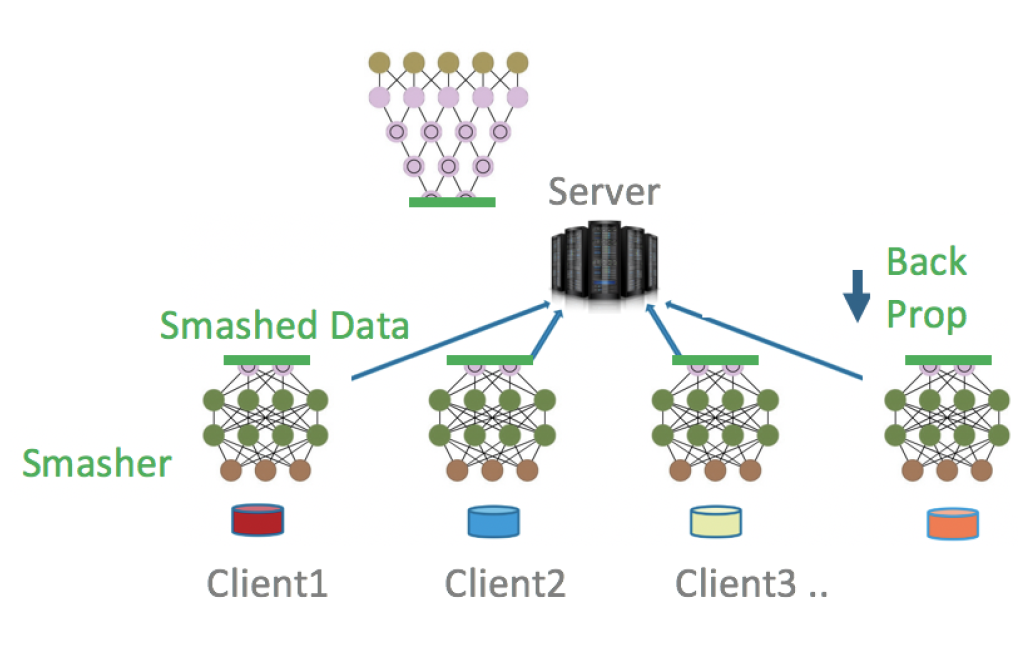


图3 分割学习多用户示意图[1]

**[注2]：对于Client#2 而言，需要获取Client#1处的Client model的模型参数。Client#2可以通过与Client#1直接进行通信获取，也可以通过Server获取，即Client#1将Client model参数上传至Server，再由Server下发至Client#2。**

如图4所示，多用户场景下的SL，训练需要使用多个Client的本地数据。当Client#1训练完成，进行Client#2的训练时，Client#2需要获取Client#1的模型参数。其获取方法可分为两类[2]：a). 中心模式; b). Peer-to-Peer（P2P）模式。如图4所示：

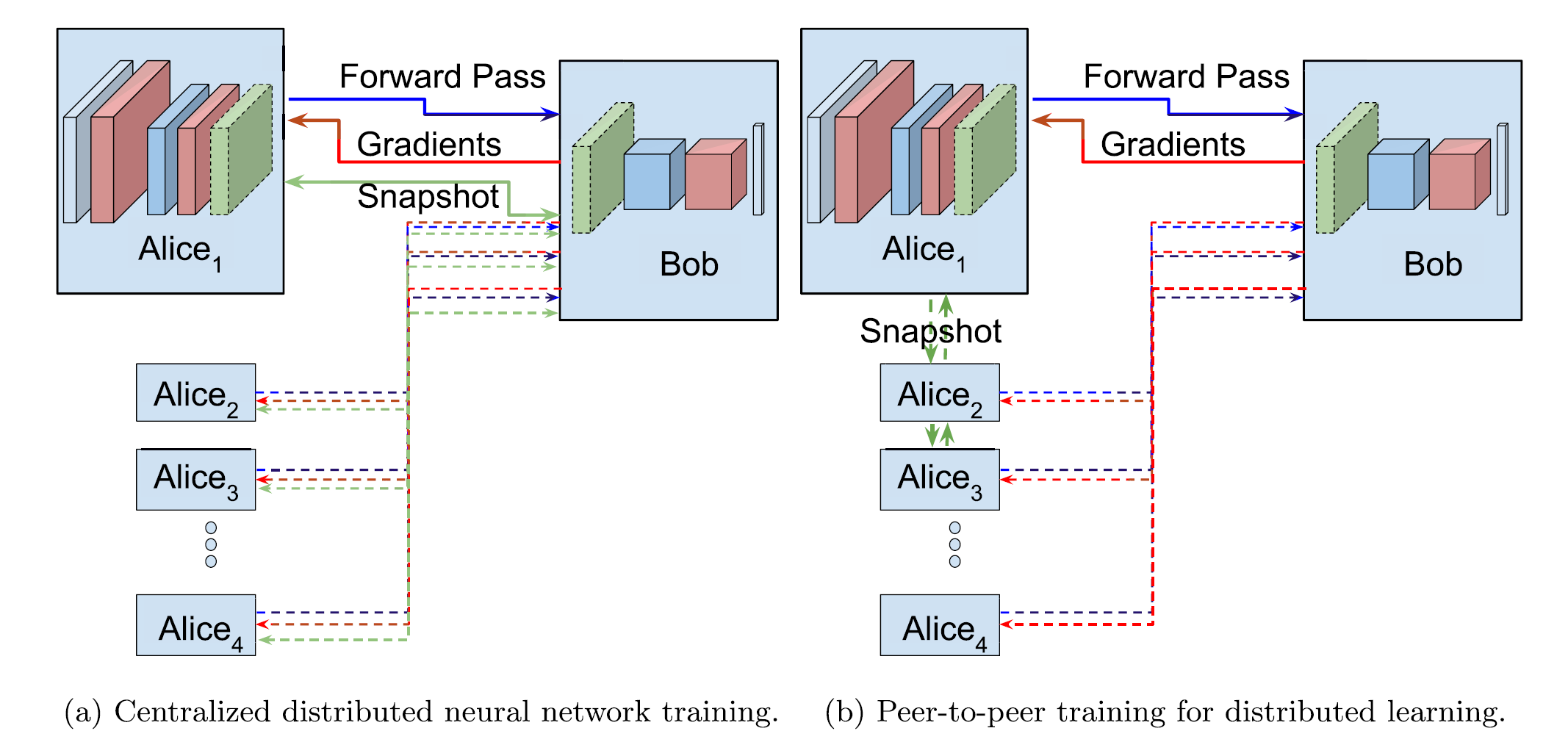


图4 分割学习的网络拓扑结构示意图 （中心模型与P2P模型）

a). 中心模式下（Client#1 --> Client#2）， Client#1(Alice1) 将参数上传至Server（Bob），Server与Client#2建立连接后将参数放送到Client#2，从而继续进行SL时；

b). P2P模式下（Client#1 --> Client#2），Server建立Client#1与Client#2的连接，Client#1直接将参数发送至Client#2。如gNB下的两个用户的sidelink，D2D参数传输。

* 1. **联邦学习与分割学习的结合 – SplitFed**

对于联邦学习和分割学习，两者的共同优势都在于：在不共享数据的前提下，完成对模型的训练（即数据可用不可见）。联邦学习上传的是训练后的模型，分割学习上传的是中间数据。

联邦学习的缺点：1). 联邦学习中，Client和Server均能获取本地和全局的模型信息。然而在分割学习中，模型在Client和Server之间进行了拆分，双方只能获取各自对应的模型部分。因此，联邦学习与分割学习相比，隐私保护性较弱； 2). 联邦学习需要Client运行整个模型，对于资源受限的终端用户而言无法承担（内存不足或运算能力不足）。而分割学习只需运行模型的一小部分即可，大大减小了运行的负载。因此，联邦学习与分割学习相比，对终端用户的要求更高。

分割学习的缺点： 每轮训练只能有一个client参加，导致训练总时间过长。

因此，文献[3]提出了一种结合FL和SL的方法，即分割联邦学习（SFL，Splitfed Learning）。其结构如图5所示，SFL的结构由三部分组成：用户（Clients）、联邦服务器（Fed Server）和主服务器（Main Server）。深度神经网络模型（）被分割成两部分：客户端模型（Client-side model）和网络端模型（Server-side model）。其中，Fed Server基于接收到的各个客户端本地模型（Client-side local model）进行聚合得到客户端全局模型（Client-side global model），并将聚合后的Client-side global model下发给各个Client。Main server负责各个Client对应的分割后的Server-side model（我们称之为Server-side local model）的训练，并且各训练可以同步进行。此外，在隐私保护允许的情况下，Fed Server的功能也可由Main Server执行。此时，网络中只存在Main Server和Clients两部分[4-5]。

其具体工作流程如下：1). Clients在client-side model上并行执行前向传播，输出中间数据并上传至Main server；2). Main Server基于接收的各个Client的中间数据在server-model上执行前向传播和反向传播（该过程针对各个Client独立且并行），更新各个client对应的server-side local model；3). Main Server将计算得到的中间数据的梯度传回对应的各个Client；4). Main Server针对server-side local model做全局聚合；5). 各个Client执行反向传播，更新其client-side local model，并上传至Fed Server; 6). Fed Server执行全局聚合并将聚合后的client-side global model下发至各Client。

注：3）和4）不分先后。

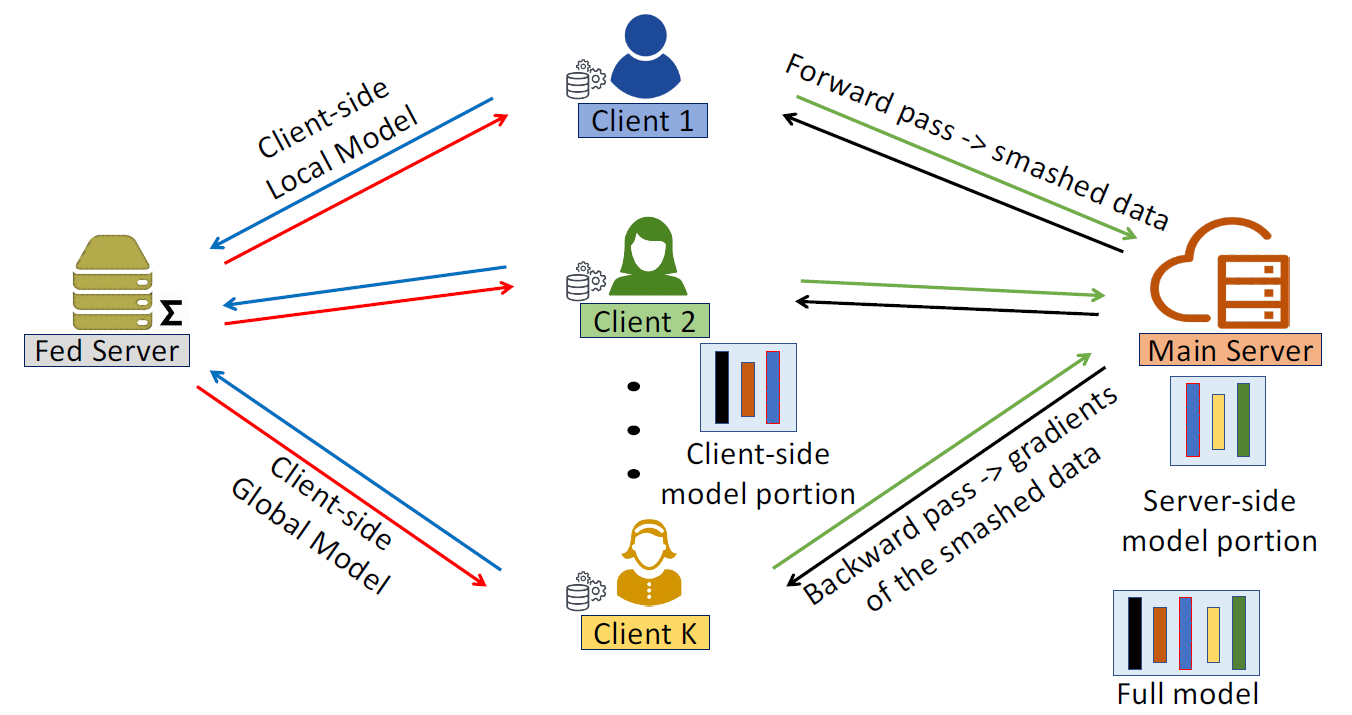


图5. 分割联邦学习结构示意图

该方法的缺点在于需要额外的Fed Server，进行Client-side model的聚合。如果将Fed Server的功能放置在Main Server，隐私又难以得到保护。

* 1. **联邦学习的D2D计算卸载**

[6] 提出了一种面向联邦学习的D2D 计算任务卸载方案，不同边缘节点通过D2D 通信**交换数据样本**，平衡节点的处理能力和任务负载，以最小化联合学习模型训练过程的总时延。

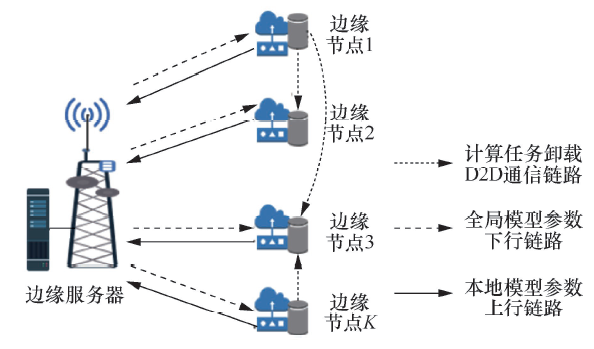


图6. 联邦学习中D2D数据卸载[6]。

该方法通过D2D通信，将数据量大、计算能力弱的边缘节点中的部分数据样本 卸载到 计算能力强、数据量小的边缘节点。实现边缘节点的计算任务量的重新分配。

1. **发明内容**

以车联网（V2X）为例，我们提出了一种新的分割学习辅助的联邦学习网络（SL-aided FL network）。如图7所示，

图片包含 图形用户界面

描述已自动生成

图7. 分割学习辅助的联邦学习（SL-aided FL）网络结构

该分布式网络下，多个**Member**与gNB或者Server一同进行联邦学习训练。其中，每个Member可以

a). 只由单个UE组成，我们称之为**Standalone Member**。此时该UE处训练完整的神经网络模型，如Member#i只由UE#i构成；

b). 由单个UE及其附近的辅助设备（Assistant Equipment，AE）共同组成，我们称之为**Non-Standalone** **Member**。此时完整的神经模型在UE和AE之间进行分割，拆分成为UE model和AE model。如Member#1由UE#1和AE#1组成，分别对应于被UE#1 model和AE#1 model。且UE#1 model和AE#1 model拼接起构成完整的神经网络模型。此外对于不同Non-Standalone Member，其UE model和AE model之间的split point可能不同，如Member#1和Member#k。

Standalone Member和Non-Standalone Member均相当于 传统的FL网络下的Client。其中，Standalone Member与传统FL下的Client没有任何差别。Non-Standalone Member则增加了AE，通过将需要训练的神经网络模型进行分割，将 一部分原本需要在UE端完成训练的模型 放置在AE端由AE完成。

在FL训练开始前: 首先，网络中的设备[注3]向gNB或者Server上报自身的状态信息（通信能力，计算能力，数据量等）。其次，gNB或Server根据各设备上传的状态信息，并结合模型的大小：1). 进行UE的选择； 2). 决定该UE作为Standalone Member/ Non-Standalone Member参与训练；3). 如果作为Non-Standalone Member，需要额外选定与其相连接的AE，以及神经网络模型的split point[注4] ；4). 为Standalone Member分配uplink传输资源，以及为Non-Standalone Member中的UE和AE分配uplink传输资源和sidelink传输资源（同时决定UE model和AE model如何上传）。最后将决定的结果通知被相应的UE和AE （Standalone Member/Non-Standalone Member，与哪个AE或者UE组成Non-Standalone Member, split point在哪一层）。

**[注3] 网络中的设备包括可作为UE或AE参与FL的设备。**

**[注4] 对于UE的选择、是否作为Standalone Member/ Non-Standalone Member、 AE的选择以及split point选择的具体准则，算法和方案，不在本专利的讨论范围之内。**

**[注5] 1)-5) 可能在训练过程中重新进行，即UE重新选择等。**

其训练流程如下：

* + - 1. gNB初始化全局模型参数，并下发全局模型
      2. 接收到全局模型的Standalone Member直接基于本地数据 进行本地模型的训练更新，完成后将更新的本地模型上传至gNB；对于Non-Standalone Member，基于UE的本地数据 进行本地模型的训练更新（即UE与AE之间 进行split learning，通过sidelink互传中间数据）。完成后将更新的本地模型上传至gNB
      3. gNB接收到所有Member上传的本地模型后进行全局模型聚合，并将聚合后的模型广播下发至各Member
      4. 重复2-4 直至模型收敛。

其中，对于Non-Standalone Member，UE与AE之间的Split Point由gNB提前决定。本地模型按照split point 分割成两部分，分别为位于UE的前层模型（UE model） 和 位于AE的后层模型（AE model）。此外，对于全局模型的下发，可以采取广播的形式。这种情况下UE和AE均拥有完整的全局模型，只需根据split point对各自的部分进行训练更新。全局模型的下发也可以采取点对点的方式，即gNB对UE和AE单独分配下行资源，发送对应的模型。

接下来，我们讨论split point对于模型训练和上传的影响。

假设神经网络有M层，我们用m表示split point的位置。其中，m=0表示模型完成卸载到AE，m=M表示模型未分割完全在UE处训练。图8给出了M=5时的一个示例。

许多不同颜色的灯光

中度可信度描述已自动生成

图8. 不同Split Point的示例（M=5）

对于Non-Standalone Member,

1). m=0, 模型完成卸载至AE。此时UE需将原始数据传输至AE，且整个训练过程中只需传输一次即可（gNB如果无额外通知）。AE完成本地训练后，可以将训练好的模型传回至UE，由UE将模型上传至gNB（对应于UE的uplink状态较好，AE的uplink状态较差的情况）。AE也可以直接将模型上传至gNB。

2). m=M, 模型完成在UE处训练，此时AE只作为Relay。

3). 0<m<M, UE和AE通过sidelink互传中间数据进行split learning。完成训练后，UE和AE可以各自上传模型至gNB，在gNB处进行拼接（如图7所示）。此外，由其中一方完成拼接后上传，如AE将AE model通过sidelink传给UE，UE完成拼接后将完整模型上传至gNB，反之亦然（UE将UE model通过sidelink传给AE，AE完成拼接后将完整模型上传至gNB）。

该网络结构的优点：

a). 大大减轻了UE的计算或通信的负担。参与FL的不同UE的计算或通信具有异构性，有些UE携带重要的数据但是计算或者通信能力不足。

b). 还可以充分利用网络中其他设备（即AE）的计算或通信资源。特别是通常gNB或者Server可以支持的参与FL的UE数目是有限的，因此在FL开始前需要进行UE选择。一部分未被选中但是依然希望参与FL（make contribution）的UE，可以作为AE参与训练。此外，AE也可以是网络中具有相当强的计算能力的设备，如mobile phone作为UE时，vehicle和UAV可作为AE，vehicle 作为UE时，Roadside Unit（RSU）可作为AE等。

c). Member与gNB之间只传输模型。对于Non-Standalone Member而言，AE处仅拥有部分模型参数，UE与AE之间传输中间数据，保证了数据的隐私和安全性。对于广播情况下，UE可以先执行一次或几次完整的本地模型更新，再将更新后的AE model发送给AE。

补充：权利要求5，6

UE与gNB之间信道的状态信息（如SINR, RSRP等）

--- 如 当上行链路的SINR小于某一阈值时，需要考虑与AE组成Non-Standalone member 内部进行 split learning，外部与其他member共同进行联邦学习。

UE自身的计算和存储能力信息（如CPU占用率，内存等）

--- 如 UE的存储空间 小于 训练模型的大小时，需要考虑与AE组成Non-Standalone member 内部进行 split learning，外部与其他member共同进行联邦学习。 或者，不被选取。

本地数据信息（如参与模型训练的样本数量，样本维度等）

--- 如 用户的**本地训练时间**由 用户的 样本数量， 样本维度，和本身的计算能力共同决定，当用户的本地训练时间大于某一阈值时，需要考虑与AE组成Non-Standalone member 内部进行 split learning，外部与其他member共同进行联邦学习。

电量，位置信息，移动性(如速度，方向，在本地停留时间)等

--- 如电量低于某一阈值时，考虑与AE组成Non-Standalone member 内部进行 split learning，外部与其他member共同进行联邦学习。

--- 如 所处位置周围无AE，或者移动性较强时，可考虑作为standalone member参与FL 或者 不被选取。

UE与周边设备的信道（sidelink）状态信息

--- 如周边设备（即AE）的sidelink均低于给定阈值，则UE只能作为standalone member参与FL 或者 不被选取

**参考文献：**

[1] Abhishek Singh, et al, "Detailed comparison of communication efficiency of split learning and federated learning," arXiv preprint arXiv:1909.09145, 2019.

[2] Otkrist Gupta and Ramesh Raskar, "Distributed learning of deep neural network over multiple agents," Journal of Network and Computer Applications, vol. 116, 2018.

[3] Thapa, Chandra, et al. "Splitfed: When federated learning meets split learning." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 36. No. 8. 2022.

[4] Wu, Di, et al. "Fedadapt: Adaptive offloading for IoT devices in federated learning." IEEE Internet of Things Journal (2022).

[5] Wu, Wen, et al. "Split Learning over Wireless Networks: Parallel Design and Resource Management." arXiv preprint arXiv:2204.08119 (2022).

[6] 蔡晓然, 莫小鹏, & 许杰. (2019). 面向联合学习的D2D计算任务卸载. 物联网学报, 3(4), 9.