



(12) 发明专利申请

(10) 申请公布号 CN 116432018 A

(43) 申请公布日 2023. 07. 14

(21) 申请号 202211228503.5

(22) 申请日 2022.10.09

(66) 本国优先权数据

202111662860.8 2021.12.30 CN

(71) 申请人 维沃移动通信有限公司

地址 523863 广东省东莞市长安镇维沃路1号

(72) 发明人 崇卫微 程思涵

(74) 专利代理机构 北京国昊天诚知识产权代理有限公司 11315

专利代理师 王思超

(51) Int. Cl.

G06F 18/214 (2023.01)

G06N 3/0464 (2023.01)

G06N 3/08 (2023.01)

权利要求书7页 说明书29页 附图10页

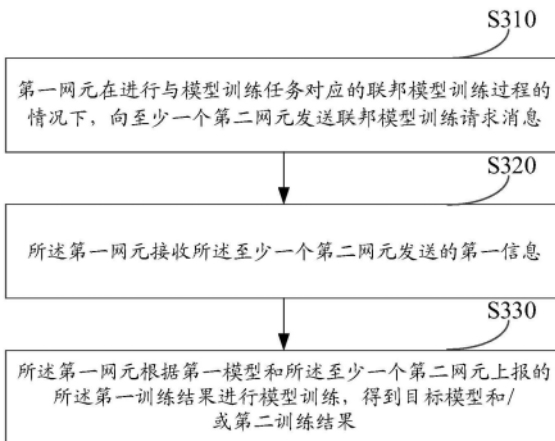
(54) 发明名称

模型训练方法、装置及通信设备

(57) 摘要

本申请公开了一种模型训练方法、装置及通信设备,属于通信技术领域,本申请实施例的模型训练方法包括:第一网元在进行与模型训练任务对应的联邦模型训练过程的情况下,向至少一个第二网元发送联邦模型训练请求消息,至少一个第二网元为参与联邦模型训练过程的网元;第一网元接收至少一个第二网元发送的第一信息,第一信息至少包括第一训练结果和/或所述第一训练结果对应的存储信息,第一训练结果对应于第二网元中用于联邦模型训练过程的训练数据;第一网元根据第一模型和至少一个第二网元上报的第一训练结果进行模型训练,得到目标模型和/或第二训练结果。

300



1. 一种模型训练方法,其特征在于,包括:

第一网元在进行与模型训练任务对应的联邦模型训练过程的情况下,向至少一个第二网元发送联邦模型训练请求消息,所述至少一个第二网元为参与所述联邦模型训练过程的网元;

所述第一网元接收所述至少一个第二网元发送的第一信息,所述第一信息至少包括第一训练结果和/或所述第一训练结果对应的存储信息,所述第一训练结果对应于所述第二网元中用于所述联邦模型训练过程的训练数据;

所述第一网元根据第一模型和所述至少一个第二网元上报的所述第一训练结果进行模型训练,得到目标模型和/或第二训练结果。

2. 如权利要求1所述的方法,其特征在于,各所述第二网元间用于所述联邦模型训练过程的训练数据对应的样本不同、但样本特征相同。

3. 如权利要求1或2所述的方法,其特征在于,所述联邦模型训练请求消息包括以下至少一项:

模型实例标识信息,所述模型实例标识信息与所述目标模型对应、且由所述第一网元分配;

所述模型训练任务的类型信息;

所述模型训练任务的标识信息;

第一指示信息,用于指示所述联邦模型训练过程是一次横向联邦学习过程;

第一过滤器的相关信息,用于限定所述模型训练任务对应的目标对象、目标时间、目标区域中的至少一项;

所述第一模型的相关信息,所述第一模型的相关信息用于各所述第二网元进行本地模型训练;

模型训练配置信息;

所述第一训练结果的上报信息;

参与所述联邦模型训练过程的各网元的相关信息。

4. 如权利要求3所述的方法,其特征在于,所述模型实例标识信息对应以下至少一项:

所述第一网元的相关信息;

第一时间,用于指示所述模型训练任务是基于所述第一时间内产生的训练数据进行;

第二时间,用于指示所述联邦模型训练过程的完成时间;

所述第二网元的相关信息。

5. 如权利要求3所述的方法,其特征在于,所述模型训练配置信息包括以下至少一项:

模型结构信息;

模型超参数信息;

所述联邦模型训练过程中的训练数据的类型信息;

模型训练条件信息,用于向所述第二网元指示所述第一网元根据所述第二网元上报的所述第一训练结果开始模型训练的条件;

模型训练次数信息,用于指示所述第二网元在向所述第一网元发送第一信息之前需要进行的本地模型训练的次数;

模型训练的时长信息,用于指示所述第二网元在向所述第一网元发送第一信息之前需

要进行的本地模型训练的时长。

6. 如权利要求5所述的方法,其特征在于,所述模型训练条件信息包括以下至少一项:

第三时间,用于指示所述第一网元对于所述第二网元反馈所述第一训练结果的等待时间达到所述第三时间时开始模型训练;

第一阈值,用于指示所述第一网元在接收到的所述第一训练结果的数量达到所述第一阈值时开始模型训练。

7. 如权利要求3所述的方法,其特征在于,所述第一训练结果的上报信息包括以下至少一项:

所述第一训练结果的上报格式;

所述第一训练结果的上报方式;

所述第一训练结果的上报条件。

8. 如权利要求1或2所述的方法,其特征在于,所述第一训练结果包括:所述第二网元根据本地训练模型进行训练得到的第二模型的模型信息和/或第二模型对应的第一梯度信息。

9. 如权利要求1或2所述的方法,其特征在于,所述第一信息还包括:

模型实例标识信息,用于所述第一网元进行模型关联。

10. 如权利要求1或2所述的方法,其特征在于,所述第二网元是所述第一网元根据所述模型训练任务从网络存储功能NRF中获取的、且能支持所述联邦模型训练过程的网元。

11. 如权利要求1或2所述的方法,其特征在于,所述第一网元包括支持所述联邦模型训练过程的服务器,所述第二网元包括支持所述联邦模型训练过程的客户端。

12. 如权利要求1-11中任一项所述的方法,其特征在于,所述方法还包括:

所述第一网元确定第一条件成立,其中,所述第一条件包括以下至少一项:

所述第一网元中没有存储或无法获取所述模型训练任务对应的全部或部分训练数据;

所述至少一个第二网元能够提供所述模型训练任务对应的全部或部分训练数据;

所述模型训练任务对应的各所述第二网元间的训练数据的样本不同、但样本特征相同。

13. 如权利要求1-11中任一项所述的方法,其特征在于,所述方法还包括:

所述第一网元在所述目标模型的损失函数的计算结果不满足预定要求或所述目标模型未收敛的情况下,向所述至少一个第二网元发送第二信息;

其中,所述第二信息至少包括所述第二训练结果和/或所述第二训练结果的存储信息,所述第二训练结果用于所述第二网元再次进行本地模型训练并重新获取第一训练结果。

14. 如权利要求13所述的方法,其特征在于,向所述至少一个第二网元发送第二信息的步骤,包括以下任一项:

对于每个所述第二网元,向所述第二网元发送指定信息,所述指定信息属于所述第二信息、且不同于所述联邦模型训练请求消息中包括的信息;

对于每个所述第二网元,向所述第二网元发送所述第二信息中的全部信息。

15. 如权利要求1-11中任一项所述的方法,其特征在于,所述方法还包括:

所述第一网元在所述目标模型的损失函数的计算结果不满足预定要求或所述目标模型未收敛的情况下,向至少一个第三网元发送第二信息;

其中,所述第二信息至少包括所述第二训练结果和/或所述第二训练结果的存储信息,所述第二训练结果用于所述第三网元进行本地模型训练以获取第三训练结果,所述第三网元是所述第一网元重新确定的、且参与所述联邦模型训练过程的网元。

16. 如权利要求15所述的方法,其特征在于,所述方法还包括:

所述第一网元向第六网元发送网元信息获取请求,所述网元信息获取请求用于请求重新确定参与联邦模型训练过程的第三网元;

所述第一网元接收所述第六网元发送的网元信息获取响应,所述网元信息获取响应中包括所述第三网元的信息。

17. 如权利要求15或16所述的方法,其特征在于,向至少一个第三网元发送第二信息之前,所述方法还包括:

所述第一网元在确定所述第二网元不再适用所述联邦模型训练过程的情况下,重新确定参与所述联邦模型训练过程的第三网元。

18. 如权利要求13或15所述的方法,其特征在于,所述第二训练结果至少包括:所述损失函数对于所述目标模型的参数的第二梯度信息或所述目标模型的信息。

19. 如权利要求13或15所述的方法,其特征在于,所述第二信息还包括以下至少一项:

模型实例标识信息,用于所述第二网元进行模型关联;

第一训练结果的上报信息;

参与所述联邦模型训练过程的各网元的相关信息;

所述目标模型对应的模型训练配置信息。

20. 如权利要求1-19中任一项所述的方法,其特征在于,所述方法还包括:

所述第一网元接收第四网元发送的模型请求消息,其中,所述模型请求消息包括以下至少一项:

所述模型训练任务的类型信息;

所述模型训练任务的标识信息;

第二过滤器的相关信息,用于限定所述模型训练任务对应的目标对象、目标时间、目标区域中的至少一项;

模型反馈相关信息,所述模型反馈相关信息包括模反馈格式、反馈条件中的至少一项。

21. 如权利要求20所述的方法,其特征在于,所述方法还包括:

所述第一网元在所述目标模型的损失函数的计算结果满足预定要求或所述目标模型收敛的情况下,向所述第四网元发送所述目标模型的相关信息;

其中,所述目标模型的相关信息包括以下至少一项:

所述模型实例标识信息;

所述目标模型信息;

第二指示信息,用于指示所述目标模型是横向联邦学习模型;

所述第二网元和/或第三网元的相关信息;

所述目标模型的存储信息。

22. 一种模型训练方法,其特征在于,所述方法包括:

第二网元接收第一网元发送的联邦模型训练请求消息,所述联邦模型训练请求消息用于请求所述第二网元参与模型训练任务对应的联邦模型训练过程;

所述第二网元根据所述联邦模型训练请求消息进行模型训练,得到第一训练结果;

所述第二网元向所述第一网元发送第一信息,所述第一信息中至少包括所述第一训练结果和/或所述第一训练结果对应的存储信息;

其中,所述第一训练结果对应于所述第二网元中用于所述联邦模型训练过程的训练数据。

23.如权利要求22所述的方法,其特征在于,所述第二网元中用于所述联邦模型训练的样本,与第五网元中用于所述联邦模型训练的样本不同、但样本特征相同,所述第五网元是参与所述联邦模型训练过程中的多个网元中除了所述第二网元之外的其他网元。

24.如权利要求22或23所述的方法,其特征在于,所述联邦模型训练请求消息包括以下至少一项:

模型实例标识信息,所述模型实例标识信息与目标模型对应、且由所述第一网元分配;

所述模型训练任务的类型信息;

所述模型训练任务的标识信息;

第一指示信息,用于指示所述联邦模型训练过程是一次横向联邦学习过程;

第一过滤器的相关信息,用于限定所述模型训练任务对应的目标对象、目标时间、目标区域中的至少一项;

第一模型的相关信息,所述第一模型的相关信息用于各所述第二网元进行本地模型训练;

模型训练配置信息;

所述第一训练结果的上报信息;

参与所述联邦模型训练过程的各网元的相关信息。

25.如权利要求24所述的方法,其特征在于,所述模型实例标识信息对应以下至少一项:

所述第一网元的相关信息;

第一时间,用于指示所述模型训练任务是基于所述第一时间内产生的训练数据进行;

第二时间,用于指示所述联邦模型训练过程的完成时间;

参与所述联邦模型训练过程的各网元的相关信息。

26.如权利要求24所述的方法,其特征在于,所述模型训练配置信息包括以下至少一项:

模型结构信息;

模型超参数信息;

所述联邦模型训练过程中的训练数据的类型信息;

模型训练条件信息,用于向所述第二网元指示所述第一网元根据第二网元上报的所述第一训练结果开始模型训练的条件;

模型训练次数信息,用于指示所述第二网元在向所述第一网元发送第一信息之前需要进行的本地模型训练的次数;

模型训练的时长信息,用于指示所述第二网元在向所述第一网元发送第一信息之前需要进行的本地模型训练的时长。

27.如权利要求26所述的方法,其特征在于,所述模型训练条件信息包括以下至少一

项：

第三时间,用于指示所述第一网元对于所述第二网元反馈所述第一训练结果的等待时间达到所述第三时间时开始模型训练;

第一阈值,用于指示所述第一网元在接收到的所述第一训练结果的数量达到所述第一阈值时开始模型训练。

28.如权利要求24所述的方法,其特征在于,所述第一训练结果的上报信息包括以下至少一项:

所述第一训练结果的上报格式;

所述第一训练结果的上报方式;

所述第一训练结果的上报条件。

29.如权利要求22或23所述的方法,其特征在于,所述第一训练结果包括:所述第二网元根据本地训练模型进行训练得到的第二模型的模型信息和/或第二模型对应的第一梯度信息。

30.如权利要求22或23所述的方法,其特征在于,所述第二网元是所述第一网元根据所述模型训练任务从网络存储功能NRF中获取的、且能支持所述联邦模型训练过程的网元。

31.如权利要求22或23所述的方法,其特征在于,所述第一网元包括支持联邦模型训练过程的服务器,所述第二网元包括支持联邦模型训练过程的客户端。

32.如权利要求22或23所述的方法,其特征在于,所述第一信息还包括以下至少一项:

模型实例标识信息,用于所述第一网元进行模型关联。

33.如权利要求24所述的方法,其特征在于,所述第二网元根据所述模型训练请求消息进行模型训练,得到第一训练结果的步骤,包括:

所述第二网元根据所述第一指示信息确定需要进行联邦模型训练过程;

所述第二网元根据所述模型训练任务的类型信息或标识信息确定所述联邦模型训练对应的模型训练任务;

所述第二网元根据所述模型训练任务以及所述第一过滤器的相关信息,获取训练数据;

所述第二网元根据所述模型训练配置信息和/或所述第一模型的相关信息,确定本地训练模型以及用于所述本地训练模型训练的模型训练方式;

所述第二网元根据所述模型训练方式对所述本地训练模型进行训练,以及基于训练后的本地训练模型对所述训练数据进行计算,得到第一训练结果。

34.如权利要求22或23所述的方法,其特征在于,所述方法还包括:

所述第二网元接收所述第一网元发送的第二信息,所述第二信息中至少包括第二训练结果和/或所述第二训练结果对应的存储信息。

35.如权利要求22或23所述的方法,其特征在于,所述第二网元接收所述第一网元发送的第二信息的步骤,包括以下任一项:

所述第二网元接收所述第一网元发送的指定信息,所述指定信息属于所述第二信息、且不同于所述联邦模型训练请求消息中包括的信息;

所述第二网元接收所述第一网元发送的第二信息中的全部信息。

36.如权利要求34所述的方法,其特征在于,所述第二信息是所述第一网元在目标模型

的损失函数的计算结果不满足预定要求或所述目标模型未收敛时发送的。

37. 如权利要求34所述的方法,其特征在于,所述第二训练结果至少包括:目标模型的损失函数对于所述目标模型的参数的第二梯度信息或目标模型的信息。

38. 如权利要求34所述的方法,其特征在于,所述第二信息还包括以下至少一项:

模型实例标识信息,用于所述第二网元进行模型关联;

目标模型的信息;

参与所述联邦模型训练过程的各网元的相关信息;

所述目标模型对应的模型训练配置信息。

39. 一种模型训练方法,其特征在于,所述方法包括:

第四网元接收第一网元发送的目标模型的相关信息;

其中,所述目标模型的相关信息至少用于表征所述目标模型为横向联邦模型。

40. 如权利要求39中所述的方法,其特征在于,所述目标模型的相关信息包括以下至少一项:

模型实例标识信息,所述模型实例标识信息与目标模型对应、且由所述第一网元分配;

第三模型信息,所述第三模型信息包括所述目标模型的网络结构信息和/或模型参数信息;

第二指示信息,用于指示所述目标模型是横向联邦学习模型;

第二网元和/或第三网元的相关信息,所述第二网元和/或所述第三网元是参与所述目标模型训练的网元;

所述目标模型的存储信息。

41. 如权利要求39中所述的方法,其特征在于,所述目标模型的相关信息是所述第一网元在损失函数的计算结果满足预定要求或所述目标模型收敛时发送的。

42. 如权利要求39中所述的方法,其特征在于,所述方法还包括:

所述第四网元向所述第一网元发送模型请求消息,其中,所述模型请求消息包括以下至少一项:

所述模型训练任务的类型信息;

所述模型训练任务的标识信息;

第二过滤器的相关信息,用于限定所述模型训练任务对应的目标对象、目标时间、目标区域中的至少一项;

模型反馈相关信息,所述模型反馈相关信息包括模型描述方式、反馈时间中的至少一项。

43. 一种模型训练方法,其特征在于,包括:

第六网元接收第一网元发送的发送网元信息获取请求,所述网元信息获取请求用于请求重新确定参与联邦模型训练过程的第三网元;

所述第六网元根据所述网元信息获取请求发送网元信息获取响应,所述网元信息获取响应中包括所述第三网元的信息。

44. 一种模型训练装置,其特征在于,包括:

第一发送模块,用于在进行与模型训练任务对应的联邦模型训练过程的情况下,向至少一个第二网元发送联邦模型训练请求消息,所述至少一个第二网元为参与所述联邦模型

训练过程的网元；

第一接收模块,用于接收所述至少一个第二网元发送的第一信息,所述第一信息至少包括第一训练结果和/或所述第一训练结果对应的存储信息,所述第一训练结果对应于所述第二网元中用于所述联邦模型训练过程的训练数据；

第一模型训练模块,用于根据第一模型和所述至少一个第二网元上报的所述第一训练结果进行模型训练,得到目标模型和/或第二训练结果。

45.一种模型训练装置,其特征在于,包括:

第二接收模块,用于第二网元接收第一网元发送的联邦模型训练请求消息,所述联邦模型训练请求消息用于请求所述第二网元参与模型训练任务对应的联邦模型训练过程；

第二模型训练模块,用于根据所述联邦模型训练请求消息进行模型训练,得到第一训练结果；

第二发送模块,用于向所述第一网元发送第一信息,所述第一信息中至少包括所述第一训练结果；

其中,所述第一训练结果对应于所述第二网元中用于所述联邦模型训练过程的训练数据。

46.一种模型训练装置,其特征在于,包括:

第三接收模块,用于接收第一网元发送的目标模型的相关信息；

其中,所述目标模型的相关信息至少用于表征所述目标模型为横向联邦模型。

47.一种通信设备,其特征在于,包括处理器和存储器,所述存储器存储可在所述处理器上运行的程序或指令,所述程序或指令被所述处理器执行时实现如权利要求1至21任一项所述的模型训练方法的步骤,或实现如权利要求22至38任一项所述的模型训练方法的步骤,或实现如权利要求39至43任一项所述的模型训练方法的步骤。

48.一种可读存储介质,其特征在于,所述可读存储介质上存储程序或指令,所述程序或指令被处理器执行时实现如权利要求1至21任一项所述的模型训练方法的步骤,或实现如权利要求22至38任一项所述的模型训练方法的步骤,或实现如权利要求39至43任一项所述的模型训练方法的步骤。

模型训练方法、装置及通信设备

技术领域

[0001] 本申请属于通信技术领域,具体涉及一种模型训练方法、装置及通信设备。

背景技术

[0002] 随着人工智能(Artificial Intelligence, AI)的快速发展,其已在各个领域得到了广泛的应用。

[0003] 其中,以通信领域为例,随着AI功能的引入,必然存在机器学习的需求。例如,对于某种智能的网络业务,其需要来自不同域、不同网元等上产生的大量数据进行模型训练,从而实现更为精准的业务实现。但是,随着数据安全和隐私问题被越来越重视,网络中的不同域、不同网元等之间存在数据隔离的问题,那么,如何基于不同域、不同网元上的数据实现模型训练成为当前急需解决的问题。

发明内容

[0004] 本申请实施例提供一种模型训练方法、装置及通信设备,能够在不进行数据共享的情况下,联合位于不同域、不同网元上的数据实现模型训练。

[0005] 第一方面,提供了一种模型训练方法,包括:第一网元在进行与模型训练任务对应的联邦模型训练过程的情况下,向至少一个第二网元发送联邦模型训练请求消息,所述至少一个第二网元为参与所述联邦模型训练过程的网元;所述第一网元接收所述至少一个第二网元发送的第一信息,所述第一信息至少包括第一训练结果和/或所述第一训练结果对应的存储信息,所述第一训练结果对应于所述第二网元中用于所述联邦模型训练过程的训练数据;所述第一网元根据第一模型和所述至少一个第二网元上报的所述第一训练结果进行模型训练,得到目标模型和/或第二训练结果。

[0006] 第二方面,提供了一种模型训练方法,所述方法包括:第二网元接收第一网元发送的联邦模型训练请求消息,所述联邦模型训练请求消息用于请求所述第二网元参与模型训练任务对应的联邦模型训练过程;所述第二网元根据所述联邦模型训练请求消息进行模型训练,得到第一训练结果;所述第二网元向所述第一网元发送第一信息,所述第一信息中至少包括所述第一训练结果;其中,所述第一训练结果对应于所述第二网元中用于所述联邦模型训练过程的训练数据。

[0007] 第三方面,提供了一种模型训练方法,所述方法包括:第四网元接收第一网元发送的目标模型的相关信息;其中,所述目标模型的相关信息至少用于表征所述目标模型为横向联邦模型。

[0008] 第四方面,提供了一种模型训练装置,包括:第一发送模块,用于在进行与模型训练任务对应的联邦模型训练过程的情况下,向至少一个第二网元发送联邦模型训练请求消息,所述至少一个第二网元为参与所述联邦模型训练过程的网元;第一接收模块,用于接收所述至少一个第二网元发送的第一信息,所述第一信息至少包括第一训练结果和/或所述第一训练结果对应的存储信息,所述第一训练结果对应于所述第二网元中用于所述联邦模

型训练过程的训练数据；第一模型训练模块，用于根据第一模型和所述至少一个第二网元上报的所述第一训练结果进行模型训练，得到目标模型和/或第二训练结果。

[0009] 第五方面，提供了一种模型训练装置，包括：第二接收模块，用于第二网元接收第一网元发送的联邦模型训练请求消息，所述联邦模型训练请求消息用于请求所述第二网元参与模型训练任务对应的联邦模型训练过程；第二模型训练模块，用于根据所述联邦模型训练请求消息进行模型训练，得到第一训练结果；第二发送模块，用于向所述第一网元发送第一信息，所述第一信息中至少包括所述第一训练结果；其中，所述第一训练结果对应于所述第二网元中用于所述联邦模型训练过程的训练数据。

[0010] 第六方面，提供了一种模型训练装置，包括：第三接收模块，用于接收第一网元发送的目标模型的相关信息；其中，所述目标模型的相关信息至少用于表征所述目标模型为横向联邦模型。

[0011] 第七方面，提供了一种通信设备，该通信设备包括处理器和存储器，所述存储器存储可在所述处理器上运行的程序或指令，所述程序或指令被所述处理器执行时实现如第一方面或第二方面或第三方面所述的方法的步骤。

[0012] 第八方面，提供了一种通信设备，包括处理器及通信接口，其中，所述通信接口和所述处理器耦合，所述处理器用于运行程序或指令，实现如第一方面所述的方法的步骤，或实现如第二方面所述的方法的步骤，或实现如第三方面所述的方法的步骤。

[0013] 第九方面，提供了一种模型训练系统，包括：第一网元、第二网元、第三网元，所述第一网元可用于执行如第一方面所述的模型训练方法的步骤，所述第二网元可用于执行如第二方面所述的模型训练方法的步骤，所述第三网元可用于执行如第三方面所述的模型训练方法的步骤。

[0014] 第十方面，提供了一种可读存储介质，所述可读存储介质上存储程序或指令，所述程序或指令被处理器执行时实现如第一方面所述的方法的步骤，或者实现如第二方面所述的方法的步骤，或者实现如第三方面所述的方法的步骤。

[0015] 第十一方面，提供了一种芯片，所述芯片包括处理器和通信接口，所述通信接口和所述处理器耦合，所述处理器用于运行程序或指令，实现如第一方面所述的方法的步骤，或实现如第二方面所述的方法的步骤，或者实现如第三方面所述的方法的步骤。

[0016] 第十二方面，提供了一种计算机程序产品/程序产品，所述计算机程序/程序产品被存储在存储介质中，所述计算机程序/程序产品被至少一个处理器执行以实现如第一方面所述的方法的步骤，或者实现如第二方面所述的方法的步骤，或者实现如第三方面所述的方法的步骤。

[0017] 在本申请实施例中，第一网元利用横向联邦模型训练的方式，联合各第二网元实现本地分布式模型训练，由此，可在不共享通信网络中各第二网元上的数据的前提下，既能确保各第二网元中的数据隐私性和数据安全性，还能确保模型训练效果。

附图说明

[0018] 图1是本申请一示例性实施例提供的无线通信系统的结构示意图。

[0019] 图2是本申请一示例性实施例提供的模型训练系统的结构示意图。

[0020] 图3是本申请实施例提供的模型训练方法的流程示意图之一。

- [0021] 图4是本申请实施例提供的模型训练方法的流程示意图之二。
- [0022] 图5是本申请实施例提供的模型训练方法的流程示意图之三。
- [0023] 图6a是本申请实施例提供的模型训练方法的交互流程示意图之一。
- [0024] 图6b是本申请实施例提供的模型训练方法的交互流程示意图之二。
- [0025] 图7是本申请实施例提供的模型训练方法的流程示意图之四。
- [0026] 图8是本申请实施例提供的模型训练方法的流程示意图之五。
- [0027] 图9是本申请实施例提供的模型训练装置的结构示意图之一。
- [0028] 图10是本申请实施例提供的模型训练装置的结构示意图之二。
- [0029] 图11是本申请实施例提供的模型训练装置的结构示意图之三。
- [0030] 图12是本申请一示例性实施例提供的通信设备的结构示意图。
- [0031] 图13是本申请一示例性实施例提供的终端的结构示意图。
- [0032] 图14是本申请实施例提供的网络侧设备的结构示意图之一。
- [0033] 图15是本申请实施例提供的网络侧设备的结构示意图之二。

具体实施方式

[0034] 下面将结合本申请实施例中的附图,对本申请实施例中的技术方案进行清楚地描述,显然,所描述的实施例是本申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员所获得的所有其他实施例,都属于本申请保护的范围。

[0035] 本申请的说明书和权利要求书中的术语“第一”、“第二”等是用于区别类似的对象,而不用于描述特定的顺序或先后次序。应该理解这样使用的术语在适当情况下可以互换,以便本申请的实施例能够以除了在这里图示或描述的那些以外的顺序实施,且“第一”、“第二”所区别的对象通常为一类,并不限定对象的个数,例如第一对象可以是一个,也可以是多个。此外,说明书以及权利要求中“和/或”表示所连接对象的至少其中之一,字符“/”一般表示前后关联对象是一种“或”的关系。

[0036] 值得指出的是,本申请实施例所描述的技术不限于长期演进型(Long Term Evolution,LTE)/LTE的演进(LTE-Advanced,LTE-A)系统,还可用于其他无线通信系统,诸如码分多址(Code Division Multiple Access,CDMA)、时分多址(Time Division Multiple Access,TDMA)、频分多址(Frequency Division Multiple Access,FDMA)、正交频分多址(Orthogonal Frequency Division Multiple Access,OFDMA)、单载波频分多址(Single-carrier Frequency-Division Multiple Access,SC-FDMA)和其他系统。本申请实施例中的术语“系统”和“网络”常被可互换地使用,所描述的技术既可用于以上提及的系统 and 无线电技术,也可用于其他系统和无线电技术。以下描述出于示例目的描述了新空口(New Radio,NR)系统,并且在以下大部分描述中使用NR术语,但是这些技术也可应用于NR系统以外的系统,如第6代(6th Generation,6G)通信系统。

[0037] 图1示出本申请实施例可应用的一种无线通信系统的框图。无线通信系统包括终端11和网络侧设备12。其中,终端11可以是手机、平板电脑(Tablet Personal Computer)、膝上型电脑(Laptop Computer)或称为笔记本电脑、个人数字助理(Personal Digital Assistant,PDA)、掌上电脑、上网本、超级移动个人计算机(ultra-mobile personal computer,UMPC)、移动上网装置(Mobile Internet Device,MID)、增强现实(augmented

reality,AR)/虚拟现实(virtual reality,VR)设备、机器人、可穿戴式设备(Wearable Device)、车载设备(VUE)、行人终端(PUE)、智能家居(具有无线通信功能的家居设备,如冰箱、电视、洗衣机或者家具等)、游戏机、个人计算机(personal computer,PC)、柜员机或者自助机等终端侧设备,可穿戴式设备包括:智能手表、智能手环、智能耳机、智能眼镜、智能首饰(智能手镯、智能手链、智能戒指、智能项链、智能脚镯、智能脚链等)、智能腕带、智能服装等。需要说明的是,在本申请实施例并不限定终端11的具体类型。网络侧设备12可以包括接入网设备或核心网设备,其中,接入网设备12也可以称为无线接入网设备、无线接入网(Radio Access Network,RAN)、无线接入网功能或无线接入网单元。接入网设备12可以包括基站、WLAN接入点或WiFi节点等,基站可被称为节点B、演进节点B(eNB)、接入点、基收发机站(Base Transceiver Station,BTS)、无线电基站、无线电收发机、基本服务集(Basic Service Set,BSS)、扩展服务集(Extended Service Set,ESS)、家用B节点、家用演进型B节点、发送接收点(Transmitting Receiving Point,TRP)或所述领域中其他某个合适的术语,只要达到相同的技术效果,所述基站不限于特定技术词汇,需要说明的是,在本申请实施例中仅以NR系统中的基站为例进行介绍,并不限定基站的具体类型。下面结合附图,通过一些实施例及其应用场景对本申请实施例提供的技术方案进行详细地说明。

[0038] 在前述无线通信系统的基础上,如图2所示,本申请实施例还提供一种模型训练系统(也可称作横向联邦学习系统或横向联邦模型训练系统)的结构示意图,该模型训练系统可以包括第一网元、第二网元、第四网元。

[0039] 所述第一网元作为所述模型训练系统中的协调者(coordinator),其可以是通信网络中某个能够提供机器学习功能的网元或设备,如网络数据分析功能(Network Data Analytics Function,NWDAF)、管理数据分析的服务(management data analytics service,MDAS)、管理数据分析的功能(management data analytic function,MDAF)等专用于提供网络智能化服务的网元或设备。或者,所述第一网元也可以是提供其他通信相关服务(如移动管理(Mobility Management,MM)服务、会话管理(Session Management,SM)服务)、且同时具备智能化功能的网元或设备,如接入和移动管理功能(Access and Mobility Management Function,AMF)、会话管理功能(Session Management Function,SMF)、应用功能(application function,AF)等,该AF可以是通信运营商部署的AF,也可以是第三方AF,又如,能够支持联邦模型训练过程的服务器等。本申请中,所述的第一网元还能够接受其他网元(如第四网元)或者设备提出的模型训练申请。

[0040] 所述第二网元作为所述模型训练系统中的训练实体,其可以是拥有本地数据但不愿意进行数据共享的网元或设备,且,该网元或设备具备本地机器学习能力。在此情况下,所述第二网元可以是但不限于无线接入网(Radio Access Network,RAN)域的AI功能网元、核心网(Core Network,CN)域的AI功能网元、第三方AI应用、UE中AI代理(client)、本地通信服务设备、支持联邦模型训练过程的客户端等等,也就是说,本申请中的第二网元可以是网络设备、终端设备或应用服务器等。当然在所述第二网元为服务器时,该服务器具备对各客户端上报的训练结果进行聚合的能力,在所述第二网元为客户端时,该客户端具备基于本地数据获取本地训练结果的能力。

[0041] 应注意,为便于理解,图2中仅示意出了两个第二网元,实际中,所述第二网元可以是一个或多个。

[0042] 所述第四网元作为模型训练系统中的模型消费者,其可以是无线通信系统中需要进行模型消费的网元或设备,如第三方AI应用、UE、网络侧设备等。可以理解,所述模型训练系统可以如图2所示包括所述第四网元(如所述联邦模型训练过程可由所述第四网元触发),也可以不包括第四网元(如所述联邦模型训练过程可由所述第一网元根据自身需求等进行触发)等。也就是,所述模型训练系统可以包括比图2中所示更多或更少的网元或设备,在此不做限制。

[0043] 下面结合附图,通过一些实施例及其应用场景对本申请实施例提供的技术方案进行详细地说明。

[0044] 如图3所示,为本申请一示例性实施例提供的模型训练方法300的流程示意图,该方法300可以但不限于由第一网元(如终端或网络侧设备)执行,具体可由安装于第一网元中的硬件和/或软件执行。本实施例中,所述方法300至少可以包括如下步骤。

[0045] S310,第一网元在进行与模型训练任务对应的联邦模型训练过程的情况下,向至少一个第二网元发送联邦模型训练请求消息。

[0046] 其中,所述模型训练任务可以是数据分析对应的AI模型训练任务、也可以是通信鲁棒性学习对应的AI模型训练任务、还可以是对UE的移动轨迹进行模型训练的AI模型训练任务等,在此不做限制。

[0047] 所述联邦模型训练过程可以由所述第一网元根据自身所需求的模型训练任务触发,也可以由第四网元触发,如第四网元可以根据自身需求向所述第一网元发送模型训练任务对应的模型请求消息,以触发所述联邦模型训练过程。

[0048] 需要注意,所述联邦模型训练过程也可以理解为横向联邦学习过程,其本质是样本的联合,适用于参与者业态相同但触达客户不同,即特征(即样本特征)重叠多、用户(即样本)重叠少的场景,如RAN域不同基站设备中不同用户或不同用户群组(如UE或UE group,即样本不同)的同一服务(如MM业务或SM业务,即样本特征相同)。基于此,本申请通过联合模型训练参与方(即第二网元)的不同样本的相同样本特征(也可称作数据特征),使训练数据对应的样本维度增多,从而训练得到一个更好的模型。

[0049] 在此情况下,所述第一网元向各所述第二网元发送的所述联邦模型训练请求消息是用于请求至少一个所述第二网元参与所述联邦模型训练过程(即所述至少一个第二网元为参与所述联邦模型训练过程的网元、且各所述第二网元间用于联邦模型训练的样本不同、但样本特征相同),如进行本地分布式模型训练,从而达到在不共享各第二网元上的原始数据(或训练数据)的前提下,联合各所述第二网元实现本地分布式模型训练,以在确保各第二网元中的数据隐私性和数据安全性的同时,确保模型训练效果。

[0050] 当然,参与所述联邦模型训练过程的第二网元的确定方式可以有很多种。一种实现方式中,所述第二网元可以是所述第一网元根据所述模型训练任务从网络存储功能(Network Repository Function,NRF)或统一数据管理实体(Unified Data Management,UDM)中获取的、且能支持所述联邦模型训练过程的网元。例如,所述第一网元从所述NRF查询获取参与此次联邦模型训练过程的AMF实例(instance(s))。又如,所述第一网元从所述UDM查询获取参与此次联邦模型训练过程的UE实例(instance(s))等,在此不做限制。

[0051] S320,所述第一网元接收所述至少一个第二网元发送的第一信息,所述第一信息至少包括第一训练结果和/或所述第一训练结果对应的存储信息。

[0052] 可以理解,各所述第二网元在接收到所述联邦模型训练请求消息后,可根据所述联邦模型训练请求消息进行模型训练,得到第一训练结果(如模型信息、模型梯度信息、模型参数变化量等),并通过第一信息将所述第一训练结果发送给所述第一网元(即所述第一信息至少包括第一训练结果和/或所述第一训练结果对应的存储信息)。其中,所述第二网元在根据所述联邦模型训练请求消息进行模型训练时,可以是基于本地预配置的初始模型实现,也可以是基于在接收所述联邦模型训练请求之前已经完成训练的本地训练模型实现,还可以基于所述联邦模型训练请求消息所指示的第一模型实现,在此不做限制。

[0053] 可选的,考虑到前述第一训练结果可以直接发送给所述第一网元,也可以存储或预存在数据库或其他设备中,以待第一网元自主获取,那么,如果所述第一训练结果是存储在数据库或其他设备中,所述第一网元可以根据所述第一训练结果的存储信息获取所述第一训练结果。可选的,在本实施例中,所述第一训练结果对应的存储信息可以包括第一训练结果的存储地址信息、存储设备标识信息、存储设备的全限定域名(Fully Qualified Domain Name,FQDN)信息等。

[0054] 需要注意的是,文中所述的第一训练结果、第二训练结果等,可以以文件或图片等方式进行表示、保存,例如包含模型、梯度、参数变化量的文件。

[0055] 此外,一种可能的实现方式中,所述第一训练结果对应于所述第二网元中用于联邦模型训练的训练数据,且各所述第二网元间用于所述联邦模型训练过程的训练数据对应的样本不同、但样本特征相同,例如,第二网元A中的训练数据对应样本S1、第二网元B中的训练数据对应样本S2、第三网元C中的训练数据对应样本S3,用于此次联邦模型训练过程的样本S1的样本特征、样本S2样本特征、样本S3样本特征均相同。S330,所述第一网元根据第一模型和所述至少一个第二网元上报的所述第一训练结果进行模型训练,得到目标模型和/或第二训练结果。

[0056] 其中,所述第一模型可以为所述第一网元为所述模型训练任务生成的一个初始模型(也可称作初始化的全局模型),或者,所述第一模型也可以是对所述第一网元对所述初始模型进行了一次或多次模型训练/更新得到的中间模型,在此不做限制。

[0057] 所述第一网元根据所述第一训练结果进行模型训练的过程在此不做限制,例如,所述第一网元可以根据所述至少一个第二网元上报的所述第一训练结果对所述第一模型进行更新等。需要注意的是,每个所述第二网元上报的第一训练结果均对应于所述第一模型,即每个所述第二网元均针对第一模型进行训练,并将训练得到的第一训练结果上报给所述第一网元,再由所述第一网元根据各所述第一训练结果对第一模型进行模型训练/更新等,以得到目标模型。可以理解,所述第一网元在基于第一模型训练得到所述目标模型时,其间可经过一次或多次联邦模型训练过程,如在所述目标模型未收敛时,可将所述目标模型作为前述的中间模型(即再次作为第一模型),继续执行联邦模型训练过程。

[0058] 此外,所述第二训练结果可以至少包括:所述目标模型的损失函数对于目标模型的参数的第二梯度信息或所述目标模型的信息(如所述目标模型的参数变化量信息等)。其中,所述目标模型的损失函数可以是所述第一网元预先为所述目标模型配置的。此外,由于所述目标模型可能仅是未收敛的中间模型,那么,对应不同训练程度(或训练次数)的目标模型,其对应的第二训练结果也可以不同,在此不做限制。本实施例中,所述第一网元利用横向联邦学习的方式,联合各第二网元实现本地分布式模型训练,由此,可在不共享通信网

络中各第二网元上的数据的前提下,既能确保各第二网元中的数据隐私性和数据安全性,还能确保模型训练效果。

[0059] 如图4所示,为本申请一示例性实施例提供的模型训练方法400的流程示意图,该方法400可以但不限于由第一网元(如终端或网络侧设备)执行,具体可由安装于第一网元中的硬件和/或软件执行。本实施例中,所述方法400至少可以包括如下步骤。

[0060] S410,第一网元在进行与模型训练任务对应的联邦模型训练过程的情况下,向至少一个第二网元发送联邦模型训练请求消息。

[0061] 其中,所述至少一个第二网元为参与所述联邦模型训练过程的网元,各所述第二网元间用于所述联邦模型训练过程的训练数据对应的样本不同、但样本特征相同。

[0062] 可以理解,S410的实现过程除了可参照方法实施例300中的相关描述之外,一种实现方式中,所述联邦模型训练请求消息至少可以包括以下(11)-(19)中的至少一项。

[0063] (11)模型实例标识信息(model instance ID)。所述模型实例标识信息用于唯一地标识所述第一网元中产生的第一模型或中间模型或目标模型(即所述模型实例标识信息与第一模型或中间模型或所述目标模型对应),以便于所述第一模型或中间模型或所述目标模型的识别、调用、消费等。

[0064] 本实施例中,所述模型实例标识信息可以由所述第一网元分配。例如,所述第一网元在确定进行与模型训练任务对应的联邦模型训练过程时,为所述模型训练任务对应分配一个模型实例标识信息,用于指示通过所述联邦模型训练过程训练得到的模型实体,即如所述第一模型或中间模型或所述目标模型。

[0065] 基于此,所述模型实例标识信息对应或包括以下(111)-(115)中的至少一项。

[0066] (111)所述第一网元的相关信息,如所述第一网元的标识信息(ID)、FQDN信息、名称信息、地址信息等。

[0067] (112)第一时间,用于指示所述模型训练任务是基于所述第一时间内产生的训练数据进行。

[0068] 可选的,所述第一时间以及后续的第二时间、第三时间可以为一个时间段或时间点,如第一时间可以为2021年10月11日等。

[0069] 此外,所述第一时间、所述第二时间、第三时间可以为相对时间或绝对时间,在此不做限制。

[0070] (113)第二时间,用于指示联邦模型训练过程的完成时间。

[0071] (114)所述第二网元的相关信息,如所述第二网元的名称信息、FQDN信息、标识信息、地址信息等。

[0072] (12)所述模型训练任务的类型信息,所述类型信息可以包括分析标识(analytics ID)、模型标识(model ID)等,以用于说明所述联邦模型训练过程是针对什么机器学习任务(即模型训练任务)而进行模型训练。

[0073] 例如,可利用字符串“analytics ID/model ID=UE mobility”来指示需要对UE的移动轨迹进行模型训练,以获取UE移动轨迹AI模型。应注意,所述类型信息除了前述字符串类型的表示方式之外,也可以利用数字表示方式或者其他编码表示方式等进行表示,此处不限。

[0074] (13)所述模型训练任务的标识信息。与前述的模型训练任务的类型信息类似,所

述模型训练任务的标识信息也可以包括analytics ID、model ID等,以用于说明所述联邦模型训练过程是针对什么机器学习任务(即模型训练任务)而进行模型训练。

[0075] 例如,可利用字符串“analytics ID/model ID=UE mobility”来指示需要对UE的移动轨迹进行模型训练,以获取UE移动轨迹AI模型。应注意,所述标识信息除了前述字符串类型的表示方式之外,也可以利用数字表示方式或者其他编码表示方式等进行表示,此处不限。

[0076] 可以理解,为避免信息冗余,所述联邦模型训练请求消息可以包括所述模型训练任务的标识信息和所述模型训练任务的类型信息中的任一种。

[0077] (14) 第一指示信息,用于指示所述联邦模型训练过程是一次横向联邦学习过程。其中,所述第一指示信息可以通过隐式或显式的方式进行指示。例如,在通过隐式方式指示时,所述第二网元可在所述联邦模型训练请求消息中包括的参与联邦模型训练过程的网元为多个这一信息,确定所述联邦模型训练过程是一次横向联邦学习过程等,在此不做限制。

[0078] (15) 第一过滤器(filter)的相关信息,其也可以理解为模型训练的过滤信息(filter information)。所述第一过滤器的相关信息用于限定所述模型训练任务对应的目标对象(如目标UE(target UE(s)))、目标时间(如目标时间周期(target time period))、目标区域(如目标感兴趣区域(target area of interest, (target AOI))中的至少一项,使得各所述第二网元可根据所述第一过滤器的相关信息采集合适的训练数据进行本地模型训练。

[0079] 例如,前述的目标对象可以是第二网元所服务的对象或对应管辖的对象,前述目标区域可以是第二网元的服务区域或对应管辖的区域,在此不做限制。

[0080] (16) 模型训练配置信息。该模型训练配置信息用于为所述横向联邦学习过程对齐模型训练方法和/或训练参数。也就是,所述模型训练配置信息用于指示所述第二网元进行本地模型训练的方法、模型算法或所采用的本地训练模型等。例如,所述第二网元可以根据所述模型训练配置信息进行本地训练模型的训练等。

[0081] 一种实现方式中,所述模型训练配置信息是由所述第一网元为所述模型训练任务确定的或所述联邦模型训练过程确定的。可选的,所述模型训练配置信息至少可以包括以下(161) - (166)中的至少一项。

[0082] (161) 模型结构信息。所述模型结构信息用于指示此次横向联邦学习训练过程使用的模型结构具体是哪种(如神经网络、深度神经网络(DNN)、卷积神经网络(CNN)、循环神经网络(RNN)、线性结构等),这使得各训练实体均使用该模型结构信息进行模型训练。

[0083] 此外,针对神经网络模型结构,所述模型结构信息中还可以细化指明神经网络的层、神经元节点、每层输入输出之间的关系等。当然,各所述第二网元也可以根据所述模型结构信息确定本次联邦模型训练过程所需使用的模型结构,在此不做限制。

[0084] (162) 模型超参数信息。所述模型超参数信息是所述目标模型外部的配置变量,一般用于模型参数的训练过程,通常可由实践者直接指定,如开始联邦模型训练过程之前设置的参数,用于定义关于模型的更高层次的概念,如复杂性或学习能力。

[0085] 可选的,所述模型超参数信息可以包括模型训练约定的算法、树的数量或树的深度、矩阵分解中潜在因素的数量、学习率、深层神经网络隐藏层数、k均值聚类中的簇数、目标模型对应的损失函数、用于判定目标模型是否收敛的预定值(与损失函数对应)等,在此

不做限制。

[0086] 需要注意,除了由所述第一网元指定所述模型超参数之外,所述第二网元也可以自己确定所述模型超参数信息。即如果所述联邦模型训练请求消息中未包括所述模型超参数信息,那么,各所述第二网元也可以自己决定所述模型超参数信息。

[0087] (163) 所述联邦模型训练过程中的训练数据的类型信息。其中,所述训练数据也可以理解为模型训练过程中的输入数据和/或输出数据,用于向第二网元推荐此次联邦模型训练过程所使用的输入数据的类型和/或输出数据的类型。其中,所述输入数据的类型可以是进行训练的样本特征等,在此不做限制。

[0088] 示例性的,在所述第二网元为两个AMF (AMF1和AMF2)时,所述训练数据的类型信息用于向两个AMF推荐此次联邦模型训练应该使用UE位置(location)、AMF load等两种输入数据进行AMF本地训练模型的训练。

[0089] (164) 模型训练条件信息,用于向所述第二网元指示所述第一网元根据所述第二网元上报的所述第一训练结果开始模型训练(或模型更新、计算等)的条件。

[0090] 一种实现方式中,所述模型训练条件信息至少可以包括以下(a)和(b)中的任意一项。

[0091] (a) 第三时间。其中,所述第三时间用于指示所述第一网元对于所述第二网元反馈所述第一训练结果的等待时间达到所述第三时间时开始模型训练。

[0092] 例如,假设参与所述联邦模型训练过程的第二网元有5个,且当所述第一网元等待5个第二网元反馈所述第一训练结果的等待时间达到所述第三时间时,所述第一网元仅接收到3个第二网元发送的第一训练结果,那么,所述第一网元可尝试基于当前所接收到的3个第二网元发送的第一训练结果开始对所述第一模型进行更新。

[0093] (b) 第一阈值。所述第一阈值用于指示所述第一网元在接收到的所述第一训练结果的数量达到所述第一阈值时开始模型训练。

[0094] 例如,假设参与所述联邦模型训练过程的第二网元有5个、第一阈值为4,那么,当所述第一网元接收到的所述第一训练结果的数量为4时,所述第一网元可尝试基于当前所接收到的4个第一训练结果开始对所述第一模型进行更新。

[0095] 又例如,假设参与所述联邦模型训练过程的第二网元有5个、第一阈值为80%,那么,当所述第一网元接收到的所述第一训练结果的数量为4,达到了80%(4/5)时,所述第一网元可尝试基于当前所接收到的4个第一训练结果开始对所述第一模型进行更新。

[0096] (165) 模型训练次数信息,用于指示所述第二网元在向所述第一网元发送第一信息之前需要进行的本地模型训练的次数。也就是,所述模型训练次数信息用于向所述第二网元指明第二网元在向第一网元返回一次第一训练结果之前需要在本地进行模型训练的本地训练次数信息。

[0097] 需要注意,所述模型训练次数信息中还可以包括此次联邦模型训练过程预计进行的训练迭代总数量(或总次数)的信息等。

[0098] (166) 模型训练的时长信息,用于指示所述第二网元在向所述第一网元发送第一信息之前需要进行的本地模型训练的时长。也就是,所述模型训练时长信息用于向所述第二网元指明第二网元在向第一网元返回一次第一训练结果之前允许在本地进行模型训练的可用时长信息。

[0099] 可以理解,前述(165)-(166)中所述的模型训练次数信息和模型训练时长信息可以是所述第二网元根据各所述第二网元的训练能力信息(如计算能力、数据存储能力、每轮训练迭代所需时长信息)确定。也就是说,对于不同的第二网元,所述模型训练次数信息和/或所述模型训练时长信息可以不同。

[0100] 例如,以所述模型训练次数为例,所述第一网元通过所述第二网元的训练能力信息确定该第二网元的模型训练次数信息,以保证训练能力弱的第二网元训练较少的次数,而训练能力强的第二网元训练较多的次数,进而使得参与联邦模型训练过程中的每个第二网元返回第一训练结果的步调同步,即,使得第一网元几乎在同时获取每个第二网元上报的第一训练结果,以进行后续模型训练和更新工作。

[0101] 另外,前述的第二网元的训练能力信息可以是所述第一网元从NRF中获取,在此不做限制。

[0102] (17) 第一训练结果的上报信息,用于限定各第二网元反馈第一训练结果时的上报格式、上报方式、上报条件等信息。

[0103] 其中,所述上报格式可以理解为模型框架信息,如基于TensorFlow、Pytorch模型框架的表达方式,或者还可以理解为基于跨框架的模型表达方式(如ONNX)等;或者,所述上报格式还可以用于指示所述第二网元返回本地训练后的整个模型、本地训练后的全部模型参数或者本地训练后发生改变的模型参数/梯度信息中的至少一种等。

[0104] 所述上报方式是用于指示直接上报第一训练结果或间接上报第一训练结果。本实施例中,直接上报第一训练结果方式用于指示第二网元将第一训练结果直接发送给第一网元;而间接上报第一训练结果用于指示第二网元需要将第一训练结果上报至特定的数据库或特定设备,在此种方式下,所述第一训练结果的上报信息中还包括第一训练结果的存储信息,如存储地址信息、存储设备标识信息、存储设备的FQDN信息等。

[0105] 所述上报条件可以包括事件触发和/或周期触发。其中,所述事件触发可以包括(171)-(173)中的任一项。

[0106] (171) 在所述第二网元训练本地模型的训练轮数(或次数)达到预定值时,所述第二网元上报第一训练结果。其中,对于不同的第二网元,第一网元指定的训练轮数可以不同,这可使得所述第一网元对齐模型训练迭代过程中各个训练实体(即第二网元)上报的第一训练结果的步调,进而确保各训练实体上报步调一致,防止出现掉队者的问题,确保模型训练效果。

[0107] (172) 在上报时间到达最长等待时间(或反馈截止时间)之前,所述第二网元上报第一训练结果。可以理解,所述最长等待时间可以理解为前述的模型训练的时长信息,即可在上报时间到达模型训练的时长信息之前,所述第二网元上报第一训练结果。

[0108] (173) 在本地模型收敛(本地模型对应的损失函数到达预设值)时,所述第二网元上报第一训练结果。

[0109] 所述周期触发是指所述第二网元可以周期性的上报第一训练结果,如每5分钟上报一次第一训练结果等。

[0110] (18) 参与所述联邦模型训练过程的各网元的相关信息,用于向第二网元指明参与此次联邦模型训练过程的其他第二网元和/或第一网元等。可选的,所述各网元的相关信息可以包括各所述第二网元的网元类型信息、标识信息或地址信息等、所述第一网元的网元

类型信息、标识信息或地址信息等。

[0111] (19)所述第一模型的相关信息,所述第一模型的相关信息用于各所述第二网元进行本地模型训练,由此,可以避免所述第二网元从零开始训练模型的过程,进而提高模型训练的效率。

[0112] 可选的,所述第一模型的相关信息可以是所述第一模型的模型标识信息、模型结构信息、模型参数信息等,在此不做限制。另外,在所述联邦模型训练过程中,在进行第一次联邦模型训练时,所述第一模型的相关信息为前述的初始模型的相关信息,在进行后续第二次、第三次……联邦模型训练时,所述第一模型的相关信息为前述的中间模型的相关信息。

[0113] 可以理解,所述联邦模型训练请求消息具体包括前述(11)-(19)中的哪个信息,其可以由协议约定、高层配置或网络侧配置,在此不做限制。此外,在参与所述联邦模型训练过程的第二网元为多个时,针对各所述第二网元的所述联邦模型训练请求消息中的内容可以部分相同,也可以部分不同,在此不做限制。

[0114] 进一步,除前述实现方式之外,在本实施例中,所述第一网元在确定是否进行与模型训练任务对应的联邦模型训练过程时,除了可以由第一网元自身或第四网元触发之外,还可能需要先确定第一条件成立或不成立,进而根据所述第一条件成立或不成立确定是否进行联邦模型训练过程,如在确定第一条件成立时,所述第一网元确定进行与模型训练任务对应的联邦模型训练过程。当然,所述第一条件也可以理解为所述第一网元确定需要进行联邦模型训练过程的因素或触发条件等。

[0115] 基于此,所述第一条件至少可以包括以下(21)-(23)中的至少一项。

[0116] (21)所述第一网元中没有存储或无法获取所述模型训练任务对应的全部或部分训练数据。例如,可能因为数据安全问题或数据隐私问题导致所述第一网元中缺乏各训练实体(即第二网元)的部分或全部训练数据,使得所述第一网元需要利用联邦模型训练过程以联合各第二网元进行本地分布式模型训练。

[0117] (22)所述至少一个第二网元能够提供所述模型训练任务对应的全部或部分训练数据。

[0118] (23)所述模型训练任务对应的各第二网元间的训练数据的样本不同、但样本特征相同。例如,为了分析A区域内(对应有多个基站管辖)所有用户的行为轨迹规律,而需要获取A区域内用户的轨迹模型,训练该轨迹模型需要的训练数据可以是不同基站下不同用户(即样本)产生的位置数据(即样本特征)。

[0119] 可以理解,所述第一条件包括前述(21)-(23)中的哪个可以由协议约定、高层配置或网络侧配置,在此不做限制。

[0120] S420,所述第一网元接收所述至少一个第二网元发送的第一信息。

[0121] 其中,所述第一信息至少包括第一训练结果和/或所述第一训练结果对应的存储信息,所述第一训练结果对应于所述第二网元中用于所述联邦模型训练过程的训练数据,即所述第一训练结果是基于所述第二网元中的训练数据得到。

[0122] 可以理解,S420的实现过程除了可参照方法实施例300中的相关描述之外,在一种可能的实现方式中,对于各所述第二网元,在接收到所述第一网元发送的联邦模型训练请求消息后,可进行模型训练得到本地训练模型,再根据所述本地训练模型计算生成所述第

一训练结果后发送给所述第一网元。

[0123] 示例性的,如果所述第二网元根据联邦模型训练请求消息进行模型训练得到本地训练模型,那么,各所述第二网元获取第一训练结果的过程可以包括如下(31)-(35),内容如下。

[0124] (31)所述第二网元根据所述第一指示信息确定需要进行联邦模型训练过程。可以理解,所述第一指示信息可从所述第一网元发送的联邦模型训练请求消息中获取,或者,所述第一指示信息可从所述第一网元直接获取,在此不做限制

[0125] (32)所述第二网元根据所述模型训练任务的类型信息或所述模型训练任务的标识信息(如analytics ID或model ID)确定所述联邦模型训练过程对应的模型训练任务。例如,所述第二网元可根据字符串“analytics ID/model ID=service experience”确定需要针对用户的某应用(application)进行模型训练以获取业务体验模型。

[0126] (33)所述第二网元根据所述模型训练任务以及所述第一过滤器的相关信息,获取训练数据。

[0127] 例如,所述第二网元可根据所述模型训练任务以及所述第一过滤器的相关信息,确定采集第二网元对应服务的小区1中在每周一07:00-09:00时段内所有UE的位置数据,作为训练数据,所采集的数据类型可以是训练实体内部配置的或者根据所述联邦模型训练请求消息中包括的所述训练数据的类型信息确定,在此不做限制。

[0128] (34)所述第二网元根据所述模型训练配置信息和/或所述第一模型的相关信息,确定本地训练模型以及用于所述本地训练模型训练的模型训练方式。

[0129] 例如,所述第二网元可根据所述模型训练配置信息确定本地训练模型的训练方式,如模型结构、算法、超参数等。又例如,所述第二网元根据所述第一模型的相关信息确定所述第一模型为所述本地训练模型。

[0130] 当然,除了根据所述第一模型的相关信息确定所述本地训练模型之外,所述本地训练模型也可以是预配置在所述第二网元中的模型,或者,所述第二网元从其他网元获取的模型,或者,所述第二网元通过本地模型训练得到的模型等,在此不做限制。

[0131] (35)所述第二网元根据所述模型训练方式对所述本地训练模型进行训练,以及基于训练后的本地训练模型对所述训练数据进行计算,得到第一训练结果。

[0132] 示例性的,假设所述第二网元为基站,那么,第*i*个基站在进行*t*轮模型训练后得到小区的模型 L_i^t ,并将该模型 L_i^t 作为第一训练结果上报给第一网元,同理,第*i*+1个基站在进行*t*轮模型训练后得到小区的模型 L_{i+1}^t ,并将模型 L_{i+1}^t 作为第一训练结果上报给所述第一网元。

[0133] 当然需要注意的是,所述第一训练结果除了可以是所述第二网元训练得到的模型(即第二模型的模型信息)之外,还可以是:所述第二网元根据本地训练模型进行训练得到的第一梯度信息、第二模型对应的参数(如结构参数、权重参数等)变化量信息等,所述第一梯度信息是所述第二网元训练得到的第二模型的损失函数对第二模型参数的梯度计算值。

[0134] 或者,在一种实现方式中,所述第一信息中还可以包括模型实例标识信息,该模型实例标识信息用于所述第一网元进行模型关联,如第一网元基于所述模型实例标识信息对不同的第二网元上报的同一第一模型得到的第一训练结果进行关联,从而针对该第一模型进行模型训练/更新。

[0135] 可以理解,所述第二网元上报的所述模型实例标识信息可根据所述联邦模型训练

请求消息中的模型实例标识信息确定。

[0136] S430,所述第一网元根据第一模型和所述至少一个第二网元上报的所述第一训练结果进行模型训练,得到目标模型和/或第二训练结果。

[0137] 可以理解,S430的实现过程除了可以参照前述方法实施例300中的相关描述之外,作为一种可能的实现方式,根据前述的模型训练条件信息的不同,所述第一网元根据第一模型和所述至少一个第二网元上报的所述第一训练结果进行模型训练,得到目标模型和/或第二训练结果的过程可以不同,下面结合方式1和方式2对所述第一网元的模型训练过程进行说明,内容如下。

[0138] 方式1:假设所有的第二网元都上报了第一训练结果或所述第一网元对于所述第二网元反馈所述第一训练结果的等待时间达到所述第三时间或所述第一网元接收到的所述第二网元反馈的第一训练结果的数量达到了所述第一阈值,那么,所述第一网元基于各个所述第二网元上报的第一训练结果对所述第一模型进行更新。

[0139] 例如,在所述第一训练结果为第二模型对应的第一梯度信息时,所述第一网元对接收到的第一梯度信息进行加权平均,获取平均模型梯度计算结果,并利用该平均模型梯度计算结果对所述第一模型进行更新。

[0140] 又例如,在所述第一训练结果为第二模型的模型信息时,假设所述第一网元接收到第*i*个第二网元上报的第二模型的模型信息为 L_i^t 、第*i*+1个第二网元上报的第二模型的模型信息为 L_{i+1}^t ,那么,所述第一网元可根据下式(1)、 L_i^t 和 L_{i+1}^t 对所述第一模型进行更新。

$$[0141] \quad G^{t+1} = G^t + \lambda \sum_{i=1}^m (L_i^t - G^t) \quad (1)$$

[0142] 其中, G^{t+1} 表示更新后的目标模型, λ 为学习率。

[0143] 方式2:所述第一网元可在每接收到一个第二网元上报的第一训练结果后,便根据该第一训练结果对所述第一模型进行更新。

[0144] 例如,假设所述第一网元接收到第*i*个第二网元上报的第一训练结果为 L_i^t ,那么,所述第一网元可根据式(2)和第一训练结果为 L_i^t 对所述第一模型进行更新。

$$[0145] \quad G^{t+1} = (1-\alpha)G^t + \alpha G^{t+1} \quad (2)$$

[0146] 其中, G^{t+1} 表示更新后的目标模型, α 为混合超参数,且 α 会随着第一模型的调参完成时间*t*和接收到的第一训练结果的时间戳 τ 的差距执行衰减策略, $\alpha_t = \alpha * s(t - \tau)$, s 为衰减系数。

[0147] 可以理解,所述第一网元在训练得到所述目标模型之后,所述第二训练结果可以是所述目标模型的损失函数对于所述目标模型的参数的梯度值,即所述第二训练结果至少包括:所述目标模型的损失函数对于所述目标模型的参数的第二梯度信息或目标模型的信息。

[0148] 当然,所述第一网元在所述目标模型的损失函数的计算结果不满足预定要求(如损失函数的精度值未达到预定值)或所述目标模型未收敛的情况下,可以再次进行联邦模型训练过程,直到所述第一网元训练得到的目标模型的损失函数的计算结果满足预定要求或所述目标模型收敛。可以理解,在所述目标模型的损失函数的计算结果不满足预定要求时,所述目标模型可以理解为中间模型,即所述第二训练结果是中间模型的损失函数对于所述中间模型的当前参数的梯度值。

[0149] 一种实现方式中,所述第一网元在再次进行联邦模型训练过程时,可再次向所述至少一个第二网元发送第二信息,所述第二信息至少包括所述第二训练结果和/或所述第二训练结果对应的存储信息(如存储地址信息、存储设备标识信息、存储设备的FQDN信息等),使得各所述第二网元基于所述第二训练结果再次进行本地训练模型的训练,以重新获取所述第一训练结果。

[0150] 例如,对应不同的第二网元,所述第一网元下发相同的第二训练结果。如,所述第一网元向RAN1(即第二网元)和RAN2分别发送目标模型的损失函数对于目标模型的当前参数的第二梯度信息。对应的,所述RAN1、所述RAN2在接收到第二训练结果后,可分别根据第二训练结果对本地模型参数进行调整,以获取调参后的本地模型,进而基于调参后的本地模型重新获取第一训练结果。

[0151] 另一种实现方式中,所述第一网元在再次进行联邦模型训练过程时,可能由于不同训练实体的数据不同,或者不同的训练实体可以有不同的数据供模型学习,又或者当训练实体掉线、存在太高的延迟等原因,不再适宜此次联邦模型训练过程,则在下一轮迭代时可换其他的训练实体,即参与联邦模型训练过程的训练实体可以与上一轮参与联邦模型训练过程的训练实体至少部分不同。

[0152] 也就是说,所述第一网元在进行下一轮联邦模型训练时,可重新确定参与联邦模型训练过程的网元,如至少一个第三网元,所述第一网元在确定参与本轮联邦模型训练的至少一个第三网元后,向至少一个第三网元发送第二信息。其中,所述第二信息至少包括所述第二训练结果和/或所述第二训练结果对应的存储信息,所述第二训练结果用于所述第三网元进行本地模型训练以获取第三训练结果。

[0153] 可以理解,前述的所述第三网元的选取可参照前述关于第二网元的选取过程的相关描述,以及所述第三训练结果的实现可参照前述第一训练结果的相关描述。示例性的,如果第一网元需要重新确定参与联邦训练过程的网元,那么,所述第一网元可以向第六网元(如NRF、UDF)发送网元信息获取请求,所述网元信息请求用于请求重新确定参与联邦模型训练过程的第三网元,所述第六网元在接收到所述网元信息获取请求的情况下,向所述第一网元发送网元信息获取响应,所述网元信息获取响应中包括所述第三网元的信息(如第三网元的地址、标识、FQDN等),那么,所述第一网元根据所述第三网元的信息确定参与联邦训练过程的第三网元。

[0154] 另外,第一网元重新确定的第三网元,是替代不适宜或不适用参与联邦模型训练过程的第二网元。也就是说,所述第一网元可以是在确定所述第二网元不再适用所述联邦模型训练过程的情况下,重新确定参与所述联邦模型训练过程的第三网元。

[0155] 需要注意的是,至少一个所述第三网元与至少一个第二网元可以至少部分相同或不同。但考虑到所述第三网元是首次参与联邦模型训练过程,因此,发送给所述第三网元的第二信息中除了包括所述目标模型信息或第二梯度信息之外,还可以包括所述联邦模型训练请求中除所述第一模型的相关信息之外的部分或全部信息,以确保所述第三网元能够顺利完成本地模型的训练。

[0156] 那么,对于所述第一网元向第二网元发送第二信息的场景,或者对于所述第三网元与所述第二网元相同、且所述第一网元向所述第三网元发送第二信息的场景,所述第一网元发送第二信息时可采用增量式发送或全量式发送。

[0157] 其中,所谓增量式发送方式是指:对于每个所述第二网元,向所述第二网元发送指定信息,所述指定信息属于所述第二信息、且不同于所述联邦模型训练请求消息中包括的信息。

[0158] 例如,如果训练实体较上一轮的迭代过程并没有改变,则所述第二信息中可以包括更新后的目标模型信息和/或第二训练结果,而不包括与所述联邦模型训练请求消息中重复的其他信息,即所述第一网元只需下发更新后的目标模型信息和/或第二训练结果而无需重复下发所述联邦模型训练请求消息中所述的其他信息,由此能够避免由于数据过大造成的资源浪费或传输可靠性差的问题。

[0159] 所谓全量式发送方式是指:对于每个所述第二网元,向所述第二网元发送所述第二信息中的全部信息。

[0160] 例如,如果训练实体较上一轮的迭代过程发生改变,则所述第二信息中除了包括更新后的目标模型信息和/或第二训练结果之外,还可以向新选择的训练实体下发所述联邦模型训练请求消息中所述的其他信息,如模型实例标识信息、第一训练结果的上报信息、模型训练配置信息、参与所述联邦模型训练过程的各网元的相关信息等。

[0161] 也就是说,除了第二训练结果之外,前述的第二信息至少还可以包括以下(51)-(55)中的至少一项。

[0162] (51)模型实例标识信息,用于所述第二网元进行模型关联。

[0163] (52)所述目标模型的信息,用于指示所述第二网元或第三网元基于该目标模型进行本地模型训练。可选的,所述目标模型的信息可以包括模型结构信息、模型超参数信息等,在此不做限制。

[0164] (53)第一训练结果的上报信息。

[0165] (54)参与所述联邦模型训练过程的各网元的相关信息。

[0166] (55)所述目标模型对应的模型训练配置信息。

[0167] 其中,关于(51)-(55)的描述可参照前述所述联邦模型训练请求消息中的相关描述,为避免重复,在此不再限制。

[0168] 需要注意,在前述联邦模型训练过程中,所述第一网元根据所述目标模型是否收敛,可以确定是否重复前述联邦模型训练过程,直至第一网元中得到的目标模型收敛。其中,所谓目标模型收敛是指所述目标模型对应的损失函数的精度值满足预算要求或达到预定值。

[0169] 如图5所示,为本申请一示例性实施例提供的模型训练方法500的流程示意图,该方法500可以但不限于由第一网元(如终端或网络侧设备)执行,具体可由安装于第一网元中的硬件和/或软件执行。本实施例中,所述方法500至少可以包括如下步骤。

[0170] S510,所述第一网元接收第四网元发送的模型请求消息。

[0171] 其中,所述模型请求消息包括以下(61)-(64)中的至少一项。

[0172] (61)所述模型训练任务的类型信息。

[0173] (62)所述模型训练任务的标识信息。

[0174] 其中,所述模型训练任务的类型信息、所述模型训练任务的标识信息可参照前述对所述联邦模型训练请求消息中的模型训练任务的类型信息、所述模型训练任务的标识信息的相关描述,为避免重复,在此不再赘述。

[0175] (63) 第二过滤器的相关信息,用于限定所述模型训练任务对应的目标对象、目标时间、目标区域中的至少一项。

[0176] (64) 模型反馈相关信息,所述模型反馈相关信息可以包括模型反馈格式、反馈条件中的至少一项。

[0177] 其中,所述模型反馈格式可以理解为模型框架信息,如基于TensorFlow、Pytorch模型框架的表达方式,或者还可以理解为基于跨框架的模型表达方式(如ONNX)等。

[0178] 所述反馈条件可以包括事件触发和/或周期触发,其中,所述事件触发包括(641)-(643)中的至少一项。

[0179] (641) 在所述第一网元训练目标模型的训练轮数(或次数)达到预定值时,所述第一网元反馈目标模型。

[0180] (642) 在训练时间到达最长等待时间之前,所述第一网元反馈目标模型。

[0181] (643) 在目标模型收敛(目标模型对应的损失函数到达预设值以及对应的预设值)时,所述第一网元反馈目标模型等。

[0182] 所述周期触发是指所述第一网元可以周期性的反馈目标模型给第四网元,如每5分钟反馈一次目标模型等。

[0183] 当然除了前述反馈条件之外,所述模型反馈信息还可以指示所述第一网元按照一次性反馈、多次持续更新反馈、反馈截止时间、周期性反馈的等方式进行所述目标模型的反馈,在此不做限制。

[0184] 需要注意,所述第一网元向各所述第二网元发送的联邦模型训练请求消息,可以是所述第一网元根据所述模型请求消息确定,例如,所述联邦模型训练请求消息中的所述模型训练任务的类型信息、所述模型训练任务的标识信息可以与所述模型请求消息中的所述模型训练任务的类型信息、所述模型训练任务的标识信息相同;又例如,所述联邦模型训练请求消息中的第一过滤器的相关信息可以与所述模型请求消息中的第二过滤器的相关信息可以相同也可以不同等,在此不做限制。

[0185] S520,第一网元在进行与模型训练任务对应的联邦模型训练过程的情况下,向至少一个第二网元发送联邦模型训练请求消息。

[0186] 其中,所述至少一个第二网元为参与所述联邦模型训练过程的网元,各所述第二网元间用于所述联邦模型训练过程的训练数据对应的样本不同、但样本特征相同;

[0187] S530,所述第一网元接收所述至少一个第二网元发送的第一信息,所述第一信息至少包括第一训练结果和/或所述第一训练结果对应的存储信息,所述第一训练结果对应于所述第二网元中用于所述联邦模型训练过程的训练数据;

[0188] S540,所述第一网元根据第一模型和所述至少一个第二网元上报的所述第一训练结果进行模型训练,得到目标模型和/或第二训练结果。

[0189] 可以理解,S520-S540的实现过程可以参照前述方法实施例300和/或400中的相关描述,为避免重复,在此不再赘述。

[0190] 当然,在一种实现方式中,所述第一网元在所述目标模型的损失函数的计算结果满足预定要求(如损失函数的精度值达到预定值,即所述目标模型收敛)或所述目标模型收敛情况下,所述第一网元可以向所述第四网元发送所述目标模型的相关信息,以用于所述第四网元基于所述目标模型执行相应的任务。

[0191] 其中,所述目标模型的相关信息包括以下(71) - (74)中的至少一项。

[0192] (71)所述模型实例标识信息。其可参照前述关于所述模型实例标识信息的描述,为避免重复,在此不再赘述。

[0193] (72)所述目标模型信息,其中所述目标模型信息包括以下(721) - (727)中的至少一项。

[0194] (721)模型结构信息。其中,所述模型结构信息用于指示所述目标模型的模型结构具体是哪一种(如神经网络、深度神经网络(DNN)、卷积神经网络(CNN)、循环神经网络(RNN)、线性结构等)。此外,针对神经网络模型结构,所述模型结构信息中还可以细化指明神经网络的层、神经元节点、每层输入输出之间的关系等。

[0195] (722)模型参数信息。其中,所述模型参数信息是所述目标模型内部的配置变量,以用于定义所述目标模型的功能,可通过数据估计或模型训练的方式得到。例如,所述模型参数可以包括人造神经网络中的权重、支持向量机中的支持向量、线性回归或逻辑回归中的系数等。

[0196] (723)模型算法信息。例如,所述模型算法信息可以但不限于包括决策树、贝叶斯分类器、K近邻、支持向量机等。

[0197] (724)模型超参数信息。其中,所述模型超参数信息是所述目标模型外部的配置变量,一般用于模型参数的训练过程,通常可由实践者直接指定。例如,所述模型超参数信息可以包括训练神经网络的学习速率、支持向量机的C和 σ 超参数、K邻域中的参数k等。

[0198] (725)模型输入数据的类型信息。

[0199] (727)模型输出数据的类型信息。

[0200] (73)第二指示信息,用于指示所述目标模型是横向联邦学习模型,或者指示该模型是经横向联邦学习过程生成的。可选的,所述第二指示信息可以显示指示或隐式指示。例如,在所述第二指示信息为隐式指示信息时,可通过模型实例标识信息或第二网元的相关信息来实现隐式指示。

[0201] (74)所述第二网元和/或所述第三网元的相关信息,如所述第二网元、第三网元的网元的类型信息、标识信息、FQDN信息、地址信息等。

[0202] 需要注意,在本实施例中,所述目标模型的相关信息除了可以包括前述(71) - (75)之外,还可以包括所述目标模型的存储信息,如存储地址信息、存储设备标识信息、FQDN等,在此不做限制。

[0203] 本实施例中,所述第一网元在接收到所述第四网元发送的模型训练请求的情况下,触发联邦模型训练过程,进而联合各所述第二网元进行本地分布式模型训练,得到目标模型,由此,能够在确保第二网元上的数据的安全性和隐私性的同时,满足第四网元上的模型训练需求。

[0204] 基于前述方法实施例200-500的描述,下面结合图6a和图6b,并以所述联邦模型训练过程是由所述第四网元触发为例,对本申请给出的模型训练方法作进一步示例性说明,内容如下。其中,假设参与联邦模型训练过程的第二网元为2个。

[0205] 示例1

[0206] 如图6a所示,示例1给出的模型训练过程如下。

[0207] S601,第四网元向所述第一网元发送模型请求消息。

- [0208] S602,所述第一网元根据所述模型请求消息确定进行联邦模型训练过程。
- [0209] S603,所述第一网元从NRF中获取能够参与联邦模型训练过程的2个第二网元,如训练实体A和训练实体B。
- [0210] S604,所述第一网元向训练实体A和训练实体B发送联邦模型训练请求消息。
- [0211] S605,训练实体A在接收到联邦模型训练请求消息后,根据所述联邦模型训练请求消息进行本地训练模型的训练,进而基于训练得到的本地训练模型获取第一训练结果。
- [0212] 相应的,训练实体B在接收到联邦模型训练请求消息后,根据所述联邦模型训练请求消息进行本地训练模型的训练,进而基于训练得到的本地训练模型获取第一训练结果。
- [0213] S606,训练实体A、训练实体B分别上报第一信息给所述第一网元,所述第一信息中至少包括所述第一训练结果。
- [0214] S607,所述第一网元基于接收到的训练实体A、训练实体B上报的第一训练结果进行模型训练,得到目标模型和/或第二训练结果。
- [0215] S608,所述第一网元确定所述目标模型对应的损失函数的计算结果是否满足预定要求或所述目标模型是否收敛。
- [0216] S609,所述第一网元在所述目标模型的损失函数的计算结果不满足预定要求或所述目标模型未收敛的情况下,向所述训练实体A、训练实体B发送第二信息,所述第二信息至少包括所述第二训练结果和/或所述第二训练结果对应的存储信息。
- [0217] S610,训练实体A、训练实体B根据接收到的第二训练结果再次进行本地模型训练,并重新获取所述第一训练结果。
- [0218] 可以理解,所述第一网元和所述训练实体A、训练实体B重复执行S605-S610,直到所述目标模型对应的损失函数的计算结果满足预定要求,即目标模型收敛,则停止所述联邦模型训练过程。
- [0219] S611,所述第一网元在所述目标模型的损失函数的计算结果满足预定要求或所述目标模型收敛的情况下,向所述第四网元发送目标模型的相关信息。
- [0220] 需要注意,前述S601-S611的实现过程可参照前述方法实施例300-500中的相关描述,并达到相同或相应的技术效果,为避免重复,在此不再赘述。
- [0221] 此外,在本实施例中,所述联邦模型训练过程可以包括但不限于前述S601-S611,例如,可以包括比前述S601-S611更多或更少的步骤,在此不做限制。
- [0222] 示例2
- [0223] 如图6b所示,示例2给出的模型训练过程如下。
- [0224] S621,第四网元向所述第一网元发送模型请求消息。
- [0225] S622,所述第一网元根据所述模型请求消息确定进行联邦模型训练过程。
- [0226] S623,所述第一网元从NRF中获取能够参与联邦模型训练过程的2个第二网元,如训练实体A和训练实体B。
- [0227] S624,所述第一网元向训练实体A和训练实体B发送联邦模型训练请求消息。
- [0228] S625,训练实体A在接收到联邦模型训练请求消息后,根据所述联邦模型训练请求消息进行本地训练模型的训练,进而基于训练得到的本地训练模型获取第一训练结果。
- [0229] S626,训练实体A上报第一信息给所述第一网元,所述第一信息中至少包括所述第一训练结果。

[0230] S627,所述第一网元基于接收到的训练实体A发送的第一训练结果进行模型训练,得到目标模型和/或第二训练结果。

[0231] S628,所述第一网元确定所述目标模型对应的损失函数的计算结果是否满足预定要求或所述目标模型是否收敛。

[0232] S629,所述第一网元在所述目标模型的损失函数的计算结果不满足预定要求或所述目标模型未收敛的情况下,向所述训练实体A发送第二信息,所述第二信息至少包括所述第二训练结果和/或所述第二训练结果对应的存储信息。

[0233] S630,训练实体A根据接收到的第二训练结果再次进行本地模型训练,以重新获取所述第一训练结果并上报,实现对联邦模型训练过程的循环迭代。

[0234] S631,训练实体B在接收到联邦模型训练请求消息后,根据所述联邦模型训练请求消息进行本地训练模型的训练,进而基于训练得到的本地训练模型获取第一训练结果。

[0235] S632,训练实体B分别上报第一信息给所述第一网元,所述第一信息中至少包括所述第一训练结果。

[0236] S633,所述第一网元基于接收到的训练实体B发送的第一训练结果进行模型训练,得到目标模型和/或第二训练结果。

[0237] S634,所述第一网元确定所述目标模型对应的损失函数的计算结果是否满足预定要求或所述目标模型是否收敛。

[0238] S635,所述第一网元在所述目标模型的损失函数的计算结果不满足预定要求或所述目标模型未收敛的情况下,向训练实体B发送第二信息,所述第二信息至少包括所述第二训练结果和/或所述第二训练结果对应的存储信息。

[0239] S636,训练实体B根据接收到的第二训练结果再次进行本地模型训练,以重新获取所述第一训练结果并上报,实现对联邦模型训练过程的循环迭代。

[0240] 可以理解,所述第一网元和所述训练实体A、训练实体B重复执行S605-S610,直到所述目标模型对应的损失函数的计算结果满足预定要求,即目标模型收敛,则停止所述联邦模型训练过程。

[0241] S637,所述第一网元在所述目标模型的损失函数的计算结果满足预定要求或所述目标模型收敛的情况下,向所述第四网元发送目标模型的相关信息。

[0242] 需要注意,前述S621-S637的实现过程可参照前述方法实施例300-500中的相关描述,并达到相同或相应的技术效果,为避免重复,在此不再赘述。

[0243] 此外,在本实施例中,所述联邦模型训练过程可以包括但不限于前述S621-S637,例如,可以包括比前述S621-S637更多或更少的步骤,在此不做限制。

[0244] 如图7所示,为本申请一示例性实施例提供的模型训练方法700的流程示意图,该方法700可以但不限于由第二网元(如终端或网络侧设备)执行,具体可由安装于第二网元中的硬件和/或软件执行。本实施例中,所述方法700至少可以包括如下步骤。

[0245] S710,第二网元接收第一网元发送的联邦模型训练请求消息,所述联邦模型训练请求消息用于请求所述第二网元参与模型训练任务对应的联邦模型训练过程。

[0246] S720,所述第二网元根据所述联邦模型训练请求消息进行模型训练,得到第一训练结果。

[0247] S730,所述第二网元向所述第一网元发送第一信息,所述第一信息中至少包括所

述第一训练结果。

[0248] 其中,所述第一训练结果对应于所述第二网元中用于所述联邦模型训练过程的训练数据,所述第二网元中用于所述联邦模型训练的样本,与第五网元中用于所述联邦模型训练的样本不同、但样本特征相同,所述第五网元是参与所述联邦模型训练过程中的多个网元中除了所述第二网元之外的其他网元。例如,假设参与模型训练的第二网元包括网元1、网元2、网元3、网元4,且执行本方法实施例700的第二网元为网元1,那么,所述第五网元则是网元2、网元3、网元4中一个或多个,在此不做限制。

[0249] 可选的,所述联邦模型训练请求消息包括以下至少一项:模型实例标识信息,所述模型实例标识信息与目标模型对应、且由所述第一网元分配;所述模型训练任务的类型信息;所述模型训练任务的标识信息;第一指示信息,用于指示所述联邦模型训练过程是一次横向联邦学习过程;第一过滤器的相关信息,用于限定所述模型训练任务对应的目标对象、目标时间、目标区域中的至少一项;第一模型的相关信息,所述第一模型的相关信息用于各所述第二网元进行本地模型训练;模型训练配置信息;所述第一训练结果的上报信息;参与所述联邦模型训练过程的各网元的相关信息。

[0250] 可选的,所述模型实例标识信息对应以下至少一项:所述第一网元的相关信息;第一时间,用于指示所述模型训练任务是基于所述第一时间内产生的训练数据进行;第二时间,用于指示所述联邦模型训练过程的完成时间;参与所述联邦模型训练过程的各网元的相关信息。

[0251] 可选的,所述模型训练配置信息包括以下至少一项:模型结构信息;模型超参数信息;所述联邦模型训练过程中的训练数据的类型信息;模型训练条件信息,用于向所述第二网元指示所述第一网元根据第二网元上报的所述第一训练结果开始模型训练的条件;模型训练次数信息,用于指示所述第二网元在向所述第一网元发送第一信息之前需要进行的本地模型训练的次数;模型训练的时长信息,用于指示所述第二网元在向所述第一网元发送第一信息之前进行的本地模型训练的时长。

[0252] 可选的,所述模型训练条件信息包括以下至少一项:第三时间,用于指示所述第一网元对于所述第二网元反馈所述第一训练结果的等待时间达到所述第三时间时开始模型训练;第一阈值,用于指示所述第一网元在接收到的所述第一训练结果的数量达到所述第一阈值时开始模型训练。

[0253] 可选的,所述第一训练结果的上报信息包括以下至少一项:所述第一训练结果的上报格式;所述第一训练结果的上报方式;所述第一训练结果的上报条件。

[0254] 可选的,所述第一训练结果包括:所述第二网元根据本地训练模型进行训练得到的第二模型的模型信息和/或第二模型对应的第一梯度信息。

[0255] 可选的,所述第二网元是所述第一网元根据所述模型训练任务从网络存储功能NRF中获取的、且能支持所述联邦模型训练过程的网元。

[0256] 可选的,所述第一信息还包括以下至少一项:模型实例标识信息,用于所述第一网元进行模型关联。

[0257] 可选的,所述第二网元根据所述模型训练请求消息进行模型训练,得到第一训练结果的步骤,包括:所述第二网元根据所述第一指示信息确定需要进行联邦模型训练过程;所述第二网元根据所述模型训练任务的类型信息或标识信息确定所述联邦模型训练对应

的模型训练任务;所述第二网元根据所述模型训练任务以及所述第一过滤器的相关信息,获取训练数据;所述第二网元根据所述模型训练配置信息和/或所述第一模型的相关信息,确定本地训练模型以及用于所述本地训练模型训练的模型训练方式;所述第二网元根据所述模型训练方式对所述本地训练模型进行训练,以及基于训练后的本地训练模型对所述训练数据进行计算,得到第一训练结果。

[0258] 可选的,所述第二网元接收所述第一网元发送的第二信息,所述第二信息中至少包括第二训练结果和/或所述第二训练结果对应的存储信息。

[0259] 可选的,所述第二网元接收所述第一网元发送的第二信息的步骤,包括以下任一项:所述第二网元接收所述第一网元发送的指定信息,所述指定信息属于所述第二信息、且不同于所述联邦模型训练请求消息中包括的信息;所述第二网元接收所述第一网元发送的第二信息中的全部信息。

[0260] 可选的,所述第二信息是所述第一网元在目标模型的损失函数的计算结果不满足预定要求或所述目标模型未收敛时发送的。

[0261] 可选的,所述第二训练结果至少包括:目标模型的损失函数对于所述目标模型的参数的第二梯度信息或目标模型的信息。

[0262] 可选的,所述第二信息还包括以下至少一项:模型实例标识信息,用于所述第二网元进行模型关联;参与所述联邦模型训练过程的各网元的相关信息;所述目标模型对应的模型训练配置信息。

[0263] 可以理解,方法实施例700中的各实现方式的实现过程可参照前述方法实施例300-600中的相关描述,并达到相同或相应的技术效果,为避免重复,在此不再赘述。

[0264] 如图8所示,为本申请一示例性实施例提供的模型训练方法800的流程示意图,该方法800可以但不限于由第四网元(如终端或网络侧设备)执行,具体可由安装于第四网元中的硬件和/或软件执行。本实施例中,所述方法800至少可以包括如下步骤。

[0265] S810,第四网元接收第一网元发送的目标模型的相关信息。

[0266] 其中,所述目标模型的相关信息至少用于表征所述目标模型为横向联邦模型。

[0267] 可选的,所述目标模型的相关信息包括以下至少一项:模型实例标识信息,所述模型实例标识信息与目标模型对应、且由所述第一网元分配;第三模型信息,所述第三模型信息包括所述目标模型的网络结构信息和/或模型参数信息;第二指示信息,用于指示所述目标模型是横向联邦学习模型;第二网元和/或第三网元的相关信息,所述第二网元和/或所述第三网元是参与所述目标模型训练的网元;所述目标模型的存储信息。

[0268] 可选的,所述目标模型的相关信息是所述第一网元在损失函数的计算结果满足预定要求或所述目标模型收敛时发送的。

[0269] 可选的,所述方法还包括:所述第四网元向所述第一网元发送模型请求消息,其中,所述模型请求消息包括以下至少一项:所述模型训练任务的类型信息;所述模型训练任务的标识信息;第二过滤器的相关信息,用于限定所述模型训练任务对应的目标对象、目标时间、目标区域中的至少一项;模型反馈相关信息,所述模型反馈相关信息包括模型描述方式、反馈时间中的至少一项。

[0270] 可以理解,方法实施例800中的各实现方式的实现过程可参照前述方法实施例300-600中的相关描述,并达到相同或相应的技术效果,为避免重复,在此不再赘述。

[0271] 本申请实施例提供的模型训练方法300-800,执行主体可以为模型训练装置。本申请实施例中以模型训练装置执行模型训练方法为例,说明本申请实施例提供的模型训练装置。

[0272] 如图9所示,为本申请一示例性实施例提供的模型训练装置900的结构示意图,所述装置900包括第一发送模块910,用于在进行与模型训练任务对应的联邦模型训练过程的情况下,向至少一个第二网元发送联邦模型训练请求消息,所述至少一个第二网元为参与所述联邦模型训练过程的网元;第一接收模块920,用于接收所述至少一个第二网元发送的第一信息,所述第一信息至少包括第一训练结果和/或所述第一训练结果对应的存储信息,所述第一训练结果对应于所述第二网元中用于所述联邦模型训练过程的训练数据;第一模型训练模块930,用于根据第一模型和所述至少一个第二网元上报的所述第一训练结果进行模型训练,得到目标模型和/或第二训练结果。

[0273] 可选的,各所述第二网元间用于所述联邦模型训练过程的训练数据对应的样本不同、但样本特征相同。

[0274] 可选的,所述联邦模型训练请求消息包括以下至少一项:模型实例标识信息,所述模型实例标识信息与所述目标模型对应、且由所述第一网元分配;所述模型训练任务的类型信息;所述模型训练任务的标识信息;第一指示信息,用于指示所述联邦模型训练过程是一次横向联邦学习过程;第一过滤器的相关信息,用于限定所述模型训练任务对应的目标对象、目标时间、目标区域中的至少一项;所述第一模型的相关信息,所述第一模型的相关信息用于各所述第二网元进行本地模型训练;模型训练配置信息;所述第一训练结果的上报信息;参与所述联邦模型训练过程的各网元的相关信息。

[0275] 可选的,所述模型实例标识信息对应以下至少一项:所述第一网元的相关信息;第一时间,用于指示所述模型训练任务是基于所述第一时间内产生的训练数据进行;第二时间,用于指示所述联邦模型训练过程的完成时间;所述第二网元的相关信息。

[0276] 可选的,所述模型训练配置信息包括以下至少一项:模型结构信息;模型超参数信息;所述联邦模型训练过程中的训练数据的类型信息;模型训练条件信息,用于向所述第二网元指示所述第一网元根据所述第二网元上报的所述第一训练结果开始模型训练的条件;模型训练次数信息,用于指示所述第二网元在向所述第一网元发送第一信息之前需要进行的本地模型训练的次数;模型训练的时长信息,用于指示所述第二网元在向所述第一网元发送第一信息之前需要进行的本地模型训练的时长。

[0277] 可选的,所述模型训练条件信息包括以下至少一项:第三时间,用于指示所述第一网元对于所述第二网元反馈所述第一训练结果的等待时间达到所述第三时间时开始模型训练;第一阈值,用于指示所述第一网元在接收到的所述第一训练结果的数量达到所述第一阈值时开始模型训练。

[0278] 可选的,所述第一训练结果的上报信息包括以下至少一项:所述第一训练结果的上报格式;所述第一训练结果的上报方式;所述第一训练结果的上报条件。

[0279] 可选的,所述第一训练结果包括:所述第二网元根据本地训练模型进行训练得到的第二模型的模型信息和/或第二模型对应的第一梯度信息。

[0280] 可选的,所述第一信息还包括以下至少一项:模型实例标识信息,用于所述第一网元进行模型关联。

[0281] 可选的,所述第二网元是所述第一网元根据所述模型训练任务从网络存储功能 NRF中获取的、且能支持所述联邦模型训练过程的网元。

[0282] 可选的,所述第一网元包括支持联邦模型训练过程的服务器,所述第二网元包括支持联邦模型训练过程的客户端。

[0283] 可选的,所述第一模型训练模块930还用于:确定第一条件成立,其中,所述第一条件包括以下至少一项:所述第一网元中没有存储或无法获取所述模型训练任务对应的全部或部分训练数据;所述至少一个第二网元能够提供所述模型训练任务对应的全部或部分训练数据;所述模型训练任务对应的各所述第二网元间的训练数据的样本不同、但样本特征相同。

[0284] 可选的,所述第一发送模块910还用于:在所述目标模型的损失函数的计算结果不满足预定要求或所述目标模型未收敛的情况下,向所述至少一个第二网元发送第二信息;其中,所述第二信息至少包括所述第二训练结果和/或所述第二训练结果对应的存储信息,所述第二训练结果用于所述第二网元再次进行本地模型训练并重新获取第一训练结果。

[0285] 可选的,所述第一发送模块910向所述至少一个第二网元发送第二信息的步骤,包括以下任一项:对于每个所述第二网元,向所述第二网元发送指定信息,所述指定信息属于所述第二信息、且不同于所述联邦模型训练请求消息中包括的信息;对于每个所述第二网元,向所述第二网元发送所述第二信息中的全部信息。

[0286] 可选的,所述第一发送模块910还用于在所述目标模型的损失函数的计算结果不满足预定要求或所述目标模型未收敛的情况下,向至少一个第三网元发送第二信息;其中,所述第二信息至少包括所述第二训练结果,所述第二训练结果用于所述第三网元进行本地模型训练以获取第三训练结果,所述第三网元是所述第一网元重新确定的、且参与所述联邦模型训练过程的网元。

[0287] 可选的,所述第一发送模块910还用于向第六网元发送网元信息获取请求,所述网元信息获取请求用于请求重新确定所述第三网元;所述第一接收模块920还用于接收所述第六网元发送的网元信息获取响应,所述网元信息获取响应中包括所述第三网元的信息。

[0288] 可选的,所述训练模块930还用于在确定所述第二网元不再适用所述联邦模型训练过程的情况下,重新确定参与所述联邦模型训练过程的第三网元。

[0289] 可选的,所述第二训练结果至少包括:所述损失函数对于所述目标模型的参数的第二梯度信息或所述目标模型的信息。

[0290] 可选的,所述第二信息还包括以下至少一项:模型实例标识信息,用于所述第二网元进行模型关联;第一训练结果的上报信息;参与所述联邦模型训练过程的各网元的相关信息;所述目标模型对应的模型训练配置信息。

[0291] 可选的,所述第一接收模块920还用于接收第四网元发送的模型请求消息,其中,所述模型请求消息包括以下至少一项:所述模型训练任务的类型信息;所述模型训练任务的标识信息;第二过滤器的相关信息,用于限定所述模型训练任务对应的目标对象、目标时间、目标区域中的至少一项;模型反馈相关信息,所述模型反馈相关信息包括模反馈格式、反馈条件中的至少一项。

[0292] 可选的,所述第一发送模块910还用于在所述目标模型的损失函数的计算结果满足预定要求或所述目标模型收敛的情况下,向所述第四网元发送所述目标模型的相关信

息;其中,所述目标模型的相关信息包括以下至少一项:所述模型实例标识信息;所述目标模型信息;第二指示信息,用于指示所述目标模型是横向联邦学习模型;所述第二网元和/或第三网元的相关信息;所述目标模型的存储信息。

[0293] 如图10所示,为本申请一示例性实施例提供的模型训练装置1000的结构示意图,所述装置1000包括:第二接收模块1010,用于第二网元接收第一网元发送的联邦模型训练请求消息,所述联邦模型训练请求消息用于请求所述第二网元参与模型训练任务对应的联邦模型训练过程;第二模型训练模块1020,用于根据所述联邦模型训练请求消息进行模型训练,得到第一训练结果;第二发送模块1030,用于向所述第一网元发送第一信息,所述第一信息中至少包括所述第一训练结果;其中,所述第一训练结果对应于所述第二网元中用于所述联邦模型训练过程的训练数据。

[0294] 可选的,所述第二网元中用于所述联邦模型训练的样本,与第五网元中用于所述联邦模型训练的样本不同、但样本特征相同,所述第五网元是参与所述联邦模型训练过程中的多个网元中除了所述第二网元之外的其他网元。

[0295] 可选的,所述联邦模型训练请求消息包括以下至少一项:模型实例标识信息,所述模型实例标识信息与目标模型对应、且由所述第一网元分配;所述模型训练任务的类型信息;所述模型训练任务的标识信息;第一指示信息,用于指示所述联邦模型训练过程是一次横向联邦学习过程;第一过滤器的相关信息,用于限定所述模型训练任务对应的目标对象、目标时间、目标区域中的至少一项;第一模型的相关信息,所述第一模型的相关信息用于各所述第二网元进行本地模型训练;模型训练配置信息;所述第一训练结果的上报信息;参与所述联邦模型训练过程的各网元的相关信息。

[0296] 可选的,所述模型实例标识信息对应以下至少一项:所述第一网元的相关信息;第一时间,用于指示所述模型训练任务是基于所述第一时间内产生的训练数据进行;第二时间,用于指示所述联邦模型训练过程的完成时间;参与所述联邦模型训练过程的各网元的相关信息。

[0297] 可选的,所述模型训练配置信息包括以下至少一项:模型结构信息;模型超参数信息;所述联邦模型训练过程中的训练数据的类型信息;模型训练条件信息,用于向所述第二网元指示所述第一网元根据第二网元上报的所述第一训练结果开始模型训练的条件;模型训练次数信息,用于指示所述第二网元在向所述第一网元发送第一信息之前需要进行的本地模型训练的次数;模型训练的时长信息,用于指示所述第二网元在向所述第一网元发送第一信息之前需要进行的本地模型训练的时长。

[0298] 可选的,所述模型训练条件信息包括以下至少一项:第三时间,用于指示所述第一网元对于所述第二网元反馈所述第一训练结果的等待时间达到所述第三时间时开始模型训练;第一阈值,用于指示所述第一网元在接收到的所述第一训练结果的数量达到所述第一阈值时开始模型训练。

[0299] 可选的,所述第一训练结果的上报信息包括以下至少一项:所述第一训练结果的上报格式;所述第一训练结果的上报方式;所述第一训练结果的上报条件。

[0300] 可选的,所述第一训练结果包括:所述第二网元根据本地训练模型进行训练得到的第的模型的模型信息和/或第二模型对应的第一梯度信息。

[0301] 可选的,所述第二网元是所述第一网元根据所述模型训练任务从网络存储功能

NRF中获取的、且能支持所述联邦模型训练过程的网元。

[0302] 可选的,所述第一网元包括支持联邦模型训练过程的服务器,所述第二网元包括支持联邦模型训练过程的客户端。

[0303] 可选的,所述第一信息还包括以下至少一项:模型实例标识信息,用于所述第一网元进行模型关联。

[0304] 可选的,所述第二模型训练模块1020根据所述模型训练请求消息进行模型训练,得到第一训练结果的步骤,包括:根据所述第一指示信息确定需要进行联邦模型训练过程;根据所述模型训练任务的类型信息或标识信息确定所述联邦模型训练对应的模型训练任务;根据所述模型训练任务以及所述第一过滤器的相关信息,获取训练数据;根据所述模型训练配置信息和/或所述第一模型的相关信息,确定本地训练模型以及用于所述本地训练模型训练的模型训练方式;根据所述模型训练方式对所述本地训练模型进行训练,以及基于训练后的本地训练模型对所述训练数据进行计算,得到第一训练结果。

[0305] 可选的,所述第二接收模块1010还用于:接收所述第一网元发送的第二信息,所述第二信息中至少包括第二训练结果和/或所述第二训练结果对应的存储信息。

[0306] 可选的,所述第二接收模块1010接收所述第一网元发送的第二信息的步骤,包括以下任一项:接收所述第一网元发送的指定信息,所述指定信息属于所述第二信息、且不同于所述联邦模型训练请求消息中包括的信息;接收所述第一网元发送的第二信息中的全部信息。

[0307] 可选的,所述第二信息是所述第一网元在目标模型的损失函数的计算结果不满足预定要求或所述目标模型未收敛时发送的。

[0308] 可选的,所述第二训练结果至少包括:目标模型的损失函数对于所述目标模型的参数的第二梯度信息或目标模型的信息;。

[0309] 可选的,所述第二信息还包括以下至少一项:模型实例标识信息,用于所述第二网元进行模型关联;参与所述联邦模型训练过程的各网元的相关信息;所述目标模型对应的模型训练配置信息。

[0310] 如图11所示,为本申请一示例性实施例提供的模型训练装置1100的结构示意图,所述装置1100包括:第三接收模块1110,用于接收第一网元发送的目标模型的相关信息;其中,所述目标模型的相关信息至少用于表征所述目标模型为横向联邦模型。

[0311] 可选的,所述目标模型的相关信息包括以下至少一项:模型实例标识信息,所述模型实例标识信息与目标模型对应、且由所述第一网元分配;第三模型信息,所述第三模型信息包括所述目标模型的网络结构信息和/或模型参数信息;第二指示信息,用于指示所述目标模型是横向联邦学习模型;第二网元和/或第三网元的相关信息,所述第二网元和/或所述第三网元是参与所述目标模型训练的网元;所述目标模型的存储信息。

[0312] 可选的,所述目标模型的相关信息是所述第一网元在损失函数的计算结果满足预定要求或所述目标模型收敛时发送的。

[0313] 可选的,所述装置1100还包括第三发送模块,用于向所述第一网元发送模型请求消息,其中,所述模型请求消息包括以下至少一项:所述模型训练任务的类型信息;所述模型训练任务的标识信息;第二过滤器的相关信息,用于限定所述模型训练任务对应的目标对象、目标时间、目标区域中的至少一项;模型反馈相关信息,所述模型反馈相关信息包括

模型描述方式、反馈时间中的至少一项。

[0314] 本申请实施例中的模型训练装置900、1000、1100可以是通信设备,例如具有操作系统的通信设备,也可以是通信设备中的部件,例如集成电路或芯片。该通信设备可以是终端,也可以为网络侧设备。示例性的,终端可以包括但不限于上述所列举的终端11的类型,网络侧设备可以包括但不限于上述所列举的网络侧设备12的类型,本申请实施例不作具体限定。

[0315] 本申请实施例提供的模型训练装置900、1000、1100能够实现图3至图8的方法实施例实现的各个过程,并达到相同的技术效果,为避免重复,这里不再赘述。

[0316] 可选的,如图12所示,本申请实施例还提供一种通信设备1200,包括处理器1201和存储器1202,存储器1202存储有可在所述处理器1201上运行的程序或指令,例如,该通信设备1200为终端时,该程序或指令被处理器1201执行时实现上述模型训练方法实施例300-800的各个步骤,且能达到相同的技术效果。该通信设备1200为网络侧设备时,该程序或指令被处理器1201执行时实现上述300-800方法实施例的各个步骤,且能达到相同的技术效果,为避免重复,这里不再赘述。

[0317] 一种实现方式中,所述通信设备1200可以为终端,该终端可以包括处理器和通信接口,所述通信接口和所述处理器耦合,所述处理器用于运行程序或指令,实现如方法实施例300-800中所述的方法的步骤。该终端实施例是与上述终端侧方法实施例对应的,上述方法实施例的各个实施过程和实现方式均可适用于该终端实施例中,且能达到相同的技术效果。具体地,图13为实现本申请实施例的一种终端的硬件结构示意图。

[0318] 该终端1300包括但不限于:射频单元1301、网络模块1302、音频输出单元1303、输入单元1304、传感器1305、显示单元1306、用户输入单元1307、接口单元1308、存储器1309、以及处理器1310等中的至少部分部件。

[0319] 本领域技术人员可以理解,终端1300还可以包括给各个部件供电的电源(比如电池),电源可以通过电源管理系统与处理器1310逻辑相连,从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。图13中示出的终端结构并不构成对终端的限定,终端可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件布置,在此不再赘述。

[0320] 应理解的是,本申请实施例中,输入单元1304可以包括图形处理单元(Graphics Processing Unit,GPU) 1041和麦克风13042,图形处理器13041对在视频捕获模式或图像捕获模式中由图像捕获装置(如摄像头)获得的静态图片或视频的图像数据进行处理。显示单元1306可包括显示面板13061,可以采用液晶显示器、有机发光二极管等形式来配置显示面板13061。用户输入单元1307包括触控面板13071以及其他输入设备13072中的至少一种。触控面板13071,也称为触摸屏。触控面板13071可包括触摸检测装置和触摸控制器两个部分。其他输入设备13072可以包括但不限于物理键盘、功能键(比如音量控制按键、开关按键等)、轨迹球、鼠标、操作杆,在此不再赘述。

[0321] 本申请实施例中,射频单元1301接收来自网络侧设备的下行数据后,可以传输给处理器1310进行处理;另外,射频单元1301可以向网络侧设备发送上行数据。通常,射频单元1301包括但不限于天线、放大器、收发信机、耦合器、低噪声放大器、双工器等。

[0322] 存储器1309可用于存储软件程序或指令以及各种数据。存储器1309可主要包括存

储程序或指令的第一存储区和存储数据的第二存储区,其中,第一存储区可存储操作系统、至少一个功能所需的应用程序或指令(比如声音播放功能、图像播放功能等)等。此外,存储器1309可以包括易失性存储器或非易失性存储器,或者,存储器1309可以包括易失性和非易失性存储器两者。其中,非易失性存储器可以是只读存储器(Read-Only Memory,ROM)、可编程只读存储器(Programmable ROM,PROM)、可擦除可编程只读存储器(Erasable PROM,EPRAM)、电可擦除可编程只读存储器(Electrically EPROM,EEPROM)或闪存。易失性存储器可以是随机存取存储器(Random Access Memory,RAM),静态随机存取存储器(Static RAM,SRAM)、动态随机存取存储器(Dynamic RAM,DRAM)、同步动态随机存取存储器(Synchronous DRAM,SDRAM)、双倍数据速率同步动态随机存取存储器(Double Data Rate SDRAM,DDRSDRAM)、增强型同步动态随机存取存储器(Enhanced SDRAM,ESDRAM)、同步连接动态随机存取存储器(Synch link DRAM,SLDRAM)和直接内存总线随机存取存储器(Direct Rambus RAM,DRRAM)。本申请实施例中的存储器1309包括但不限于这些和任意其它适合类型的存储器。

[0323] 处理器1310可包括一个或多个处理单元;可选的,处理器1310集成应用处理器和调制解调处理器,其中,应用处理器主要处理涉及操作系统、用户界面和应用程序等的操作,调制解调处理器主要处理无线通信信号,如基带处理器。可以理解的是,上述调制解调处理器也可以不集成到处理器1310中。

[0324] 其中,射频单元1301,用于在进行与模型训练任务对应的联邦模型训练过程的情况下,向至少一个第二网元发送联邦模型训练请求消息,所述至少一个第二网元为参与所述联邦模型训练过程的网元,各所述第二网元间用于所述联邦模型训练过程的训练数据对应的样本不同、但样本特征相同;以及接收所述至少一个第二网元发送的第一信息,所述第一信息至少包括第一训练结果和/或所述第一训练结果对应的存储信息,所述第一训练结果对应于所述第二网元中用于所述联邦模型训练过程的训练数据;所述处理器1310用于根据第一模型和所述至少一个第二网元上报的所述第一训练结果进行模型训练,得到目标模型和/或第二训练结果。

[0325] 本实施例中,通过利用横向联邦学习的方式,联合各第二网元实现本地分布式模型训练,由此,可在不共享通信网络中各第二网元上的数据的前提下,既能确保各第二网元中的数据隐私性和数据安全性,还能确保模型训练效果。

[0326] 另一种实现方式中,所述通信设备1200还可以为网络侧设备,所述网络侧设备包括处理器和通信接口,所述通信接口和所述处理器耦合,所述处理器用于运行程序或指令,实现如实施例300-800中所述的方法的步骤。该网络侧设备实施例是与上述网络侧设备方法实施例对应的,上述方法实施例的各个实施过程和实现方式均可适用于该网络侧设备实施例中,且能达到相同的技术效果。

[0327] 示例性的,如图14所示,为本申请实施例提供的一种网络侧设备的结构示意图。该网络侧设备1400包括:天线1401、射频装置1402、基带装置1403、处理器1404和存储器1405。天线1401与射频装置1402连接。在上行方向上,射频装置1402通过天线1401接收信息,将接收的信息发送给基带装置1403进行处理。在下行方向上,基带装置1403对要发送的信息进行处理,并发送给射频装置1402,射频装置1402对收到的信息进行处理后经过天线1401发送出去。

[0328] 以上实施例网络侧设备执行的方法可以在基带装置1403中实现,该基带装置1403包基带处理器。

[0329] 基带装置1403例如可以包括至少一个基带板,该基带板上设置有多个芯片,如图14所示,其中一个芯片例如为基带处理器,通过总线接口与存储器1405连接,以调用存储器1405中的程序,执行以上方法实施例中所示的网络设备操作。

[0330] 该网络侧设备还可以包括网络接口1406,该接口例如为通用公共无线接口(common public radio interface,CPRI)。

[0331] 具体地,本发明实施例的网络侧设备1400还包括:存储在存储器1405上并可在处理器1404上运行的指令或程序,处理器1404调用存储器1405中的指令或程序执行图9-图11所示各模块执行的方法,并达到相同的技术效果,为避免重复,故不在此赘述。

[0332] 示例性的,如图15所示,为本申请另一示例实施例提供的网络侧设备1500的结构示意图,该网络侧设备1500包括:处理器1501、网络接口1502和存储器1503。其中,网络接口1502例如为通用公共无线接口(common public radio interface,CPRI)。

[0333] 具体地,本发明实施例的网络侧设备1500还包括:存储在存储器1503上并可在处理器1501上运行的指令或程序,处理器1501调用存储器1503中的指令或程序执行图9-图11所示各模块执行的方法,并达到相同的技术效果,为避免重复,故不在此赘述。

[0334] 本申请实施例还提供一种可读存储介质,所述可读存储介质上存储有程序或指令,该程序或指令被处理器执行时实现上述模型训练方法实施例的各个过程,且能达到相同的技术效果,为避免重复,这里不再赘述。

[0335] 其中,所述处理器为上述实施例中所述的终端中的处理器。所述可读存储介质,包括计算机可读存储介质,如计算机只读存储器ROM、随机存取存储器RAM、磁碟或者光盘等。

[0336] 本申请实施例另提供了一种芯片,所述芯片包括处理器和通信接口,所述通信接口和所述处理器耦合,所述处理器用于运行网络侧设备程序或指令,实现上述模型训练方法实施例的各个过程,且能达到相同的技术效果,为避免重复,这里不再赘述。

[0337] 应理解,本申请实施例提到的芯片还可以称为系统级芯片,系统芯片,芯片系统或片上系统芯片等。

[0338] 本申请实施例还提供了一种计算机程序产品,该计算机程序产品包括处理器、存储器及存储在所述存储器上并可在所述处理器上运行的程序或指令,所述程序或指令被所述处理器执行时,实现上述模型训练方法实施例的各个过程,且能达到相同的技术效果,为避免重复,这里不再赘述。

[0339] 本申请实施例还提供了一种模型训练系统,包括:第一网元、第二网元、第三网元,所述第一网元可用于执行方法实施例300-600中的方法的步骤,所述第二网元可用于执行方法实施例700中的方法的步骤,所述第三网元可用于执行方法实施例800中的方法的步骤。

[0340] 需要说明的是,在本文中,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、物品或者装置不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、物品或者装置所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括该要素的过程、方法、物品或者装置中还存在另外的相同要素。此外,需要指出的是,本申请实

施方式中的方法和装置的范围不限按示出或讨论的顺序来执行功能,还可包括根据所涉及的功能按基本同时的方式或按相反的顺序来执行功能,例如,可以按不同于所描述的次序来执行所描述的方法,并且还可以添加、省去、或组合各种步骤。另外,参照某些示例所描述的特征可在其他示例中被组合。

[0341] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以计算机软件产品的形式体现出来,该计算机软件产品存储在一个存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端(可以是手机,计算机,服务器,空调器,或者网络设备等)执行本申请各个实施例所述的方法。

[0342] 上面结合附图对本申请的实施例进行了描述,但是本申请并不局限于上述的具体实施方式,上述的具体实施方式仅仅是示意性的,而不是限制性的,本领域的普通技术人员在本申请的启示下,在不脱离本申请宗旨和权利要求所保护的范围情况下,还可做出很多形式,均属于本申请的保护之内。

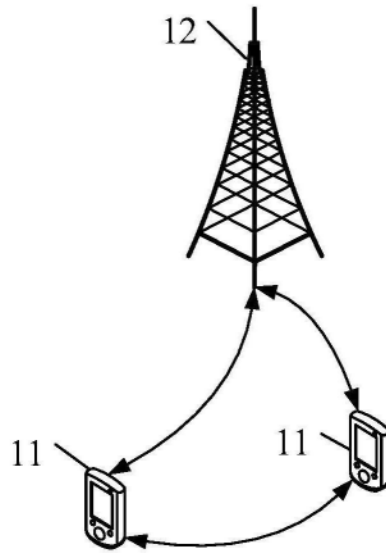


图1

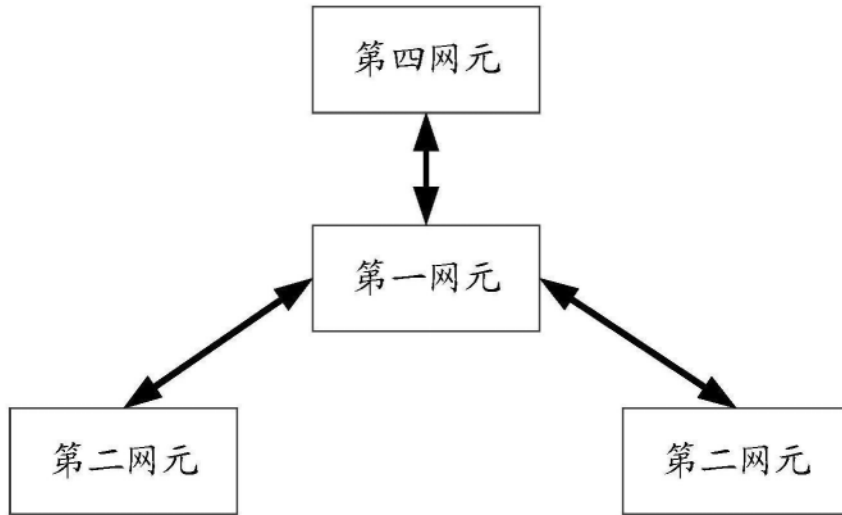


图2

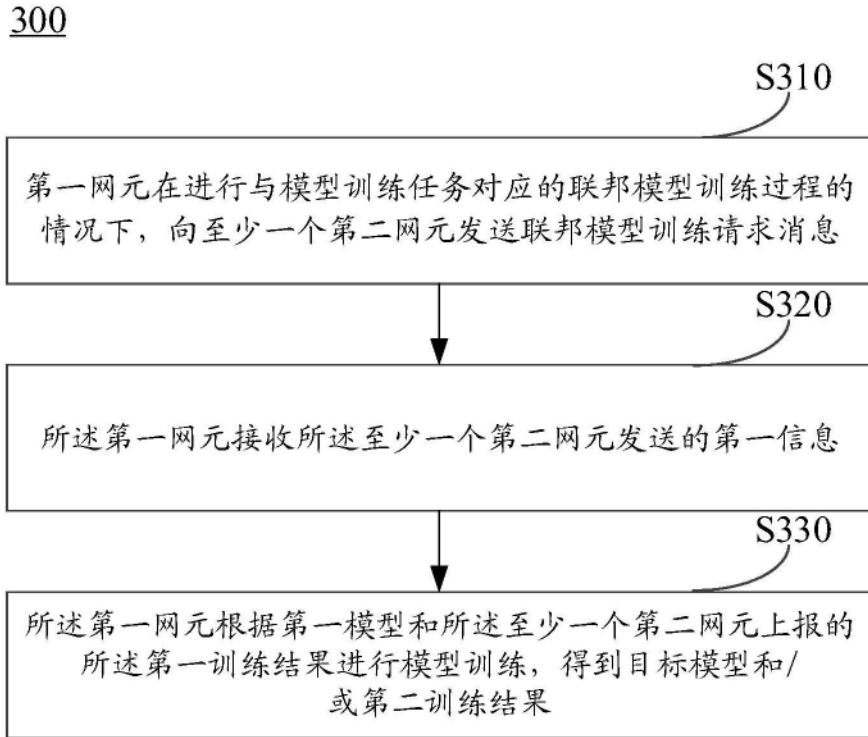


图3

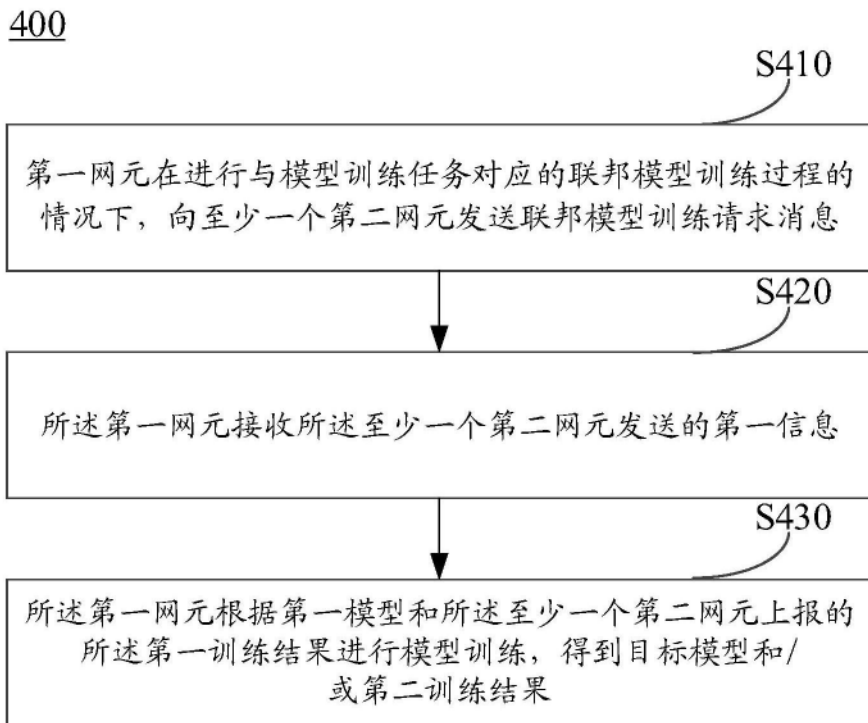


图4

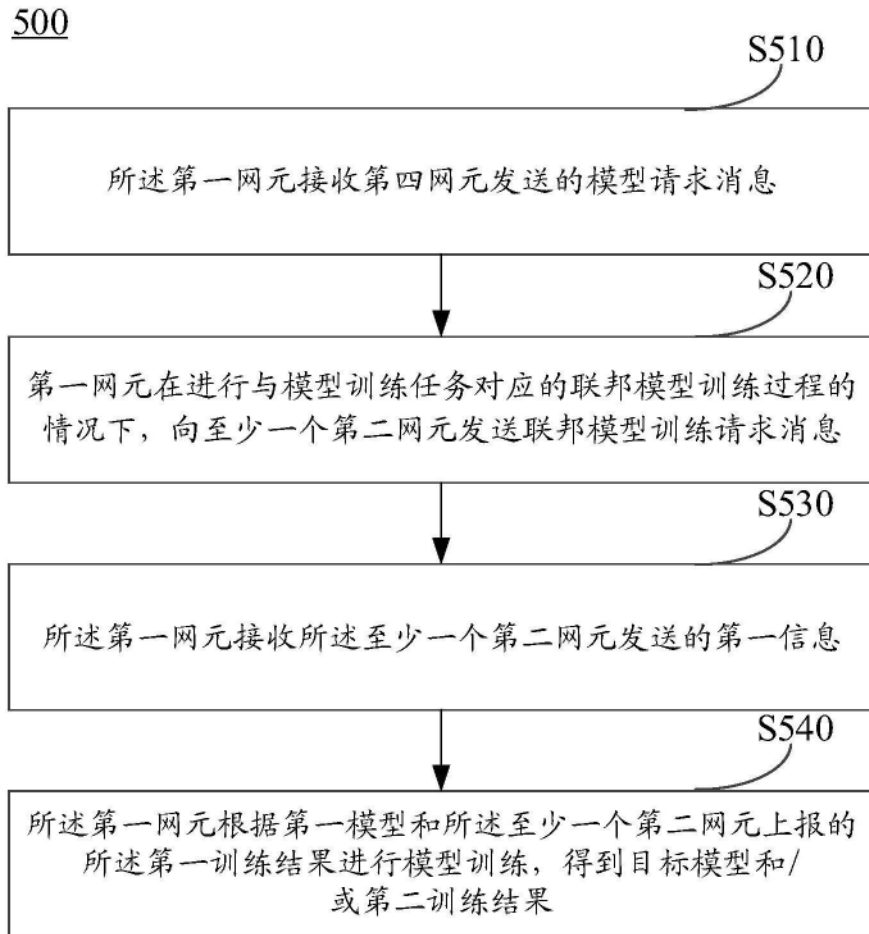


图5

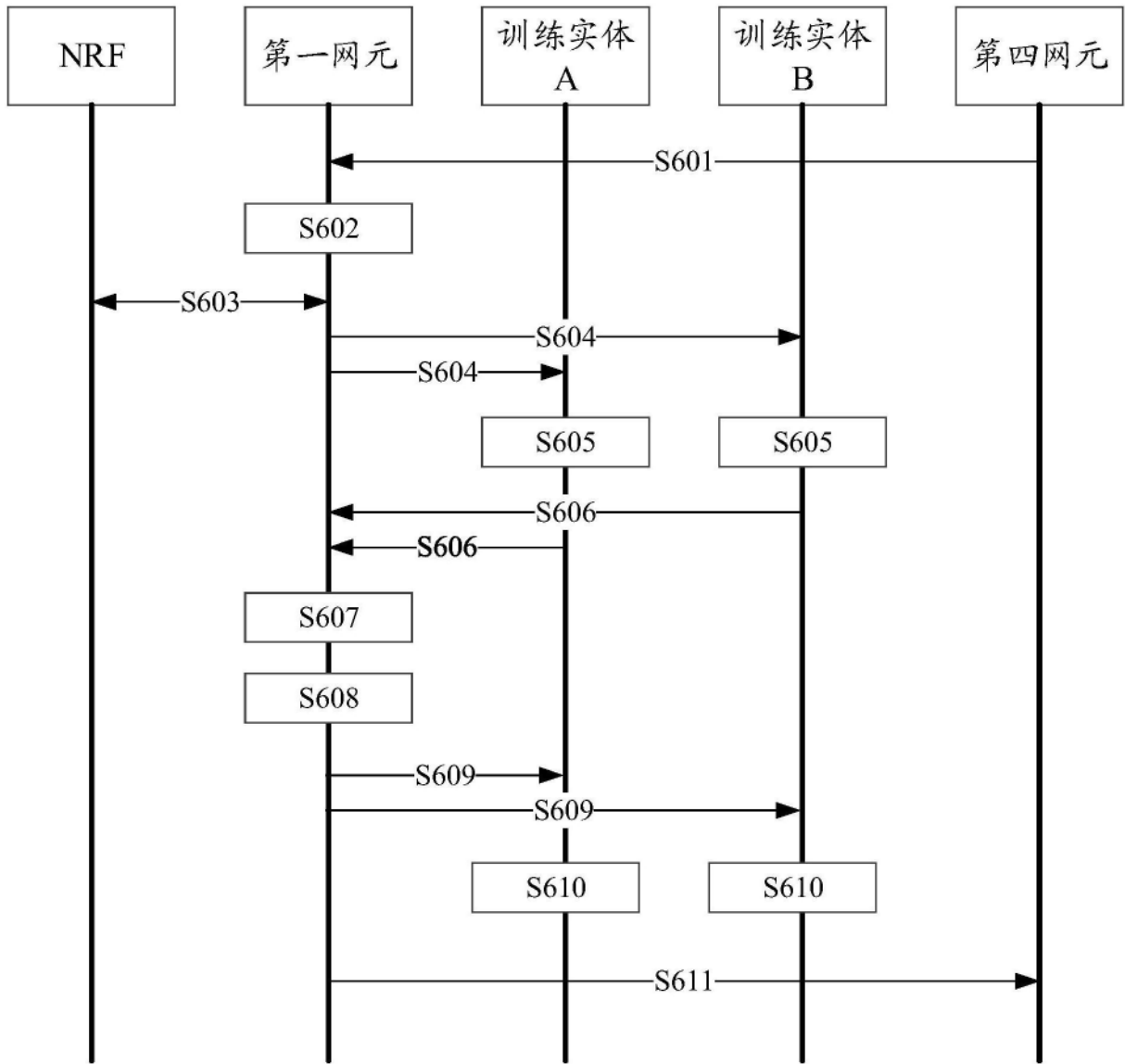


图6a

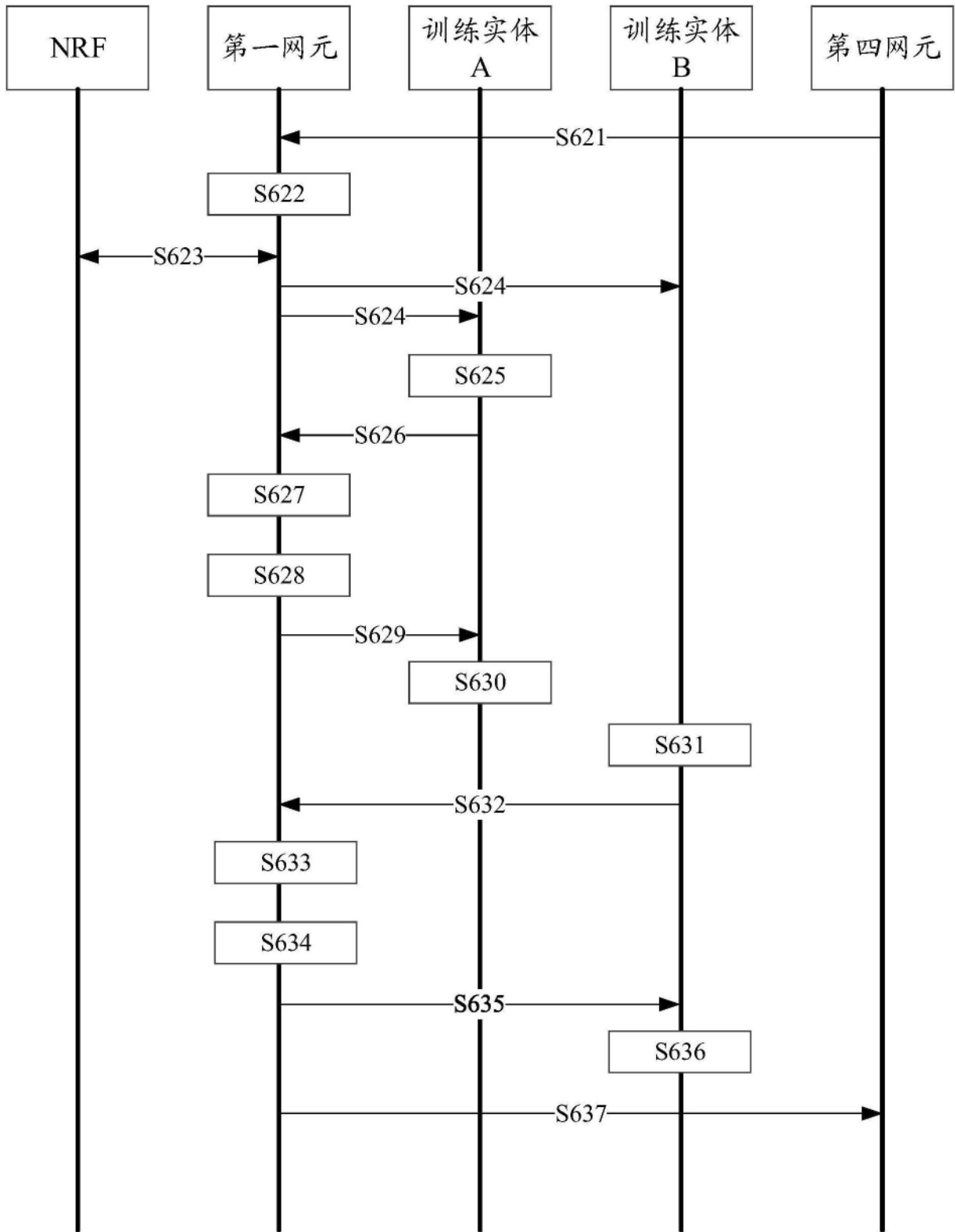


图6b

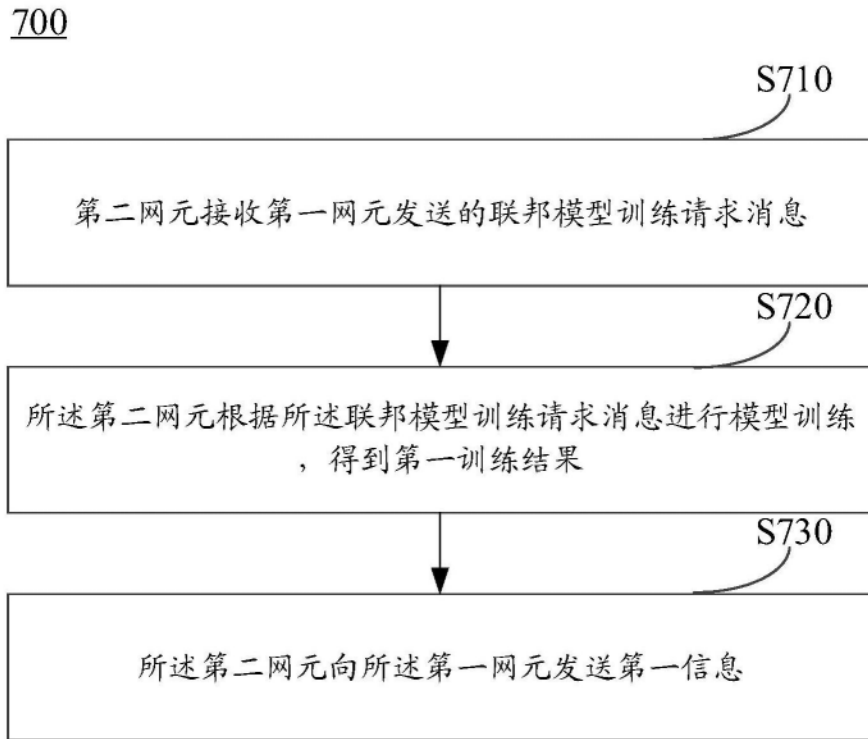


图7

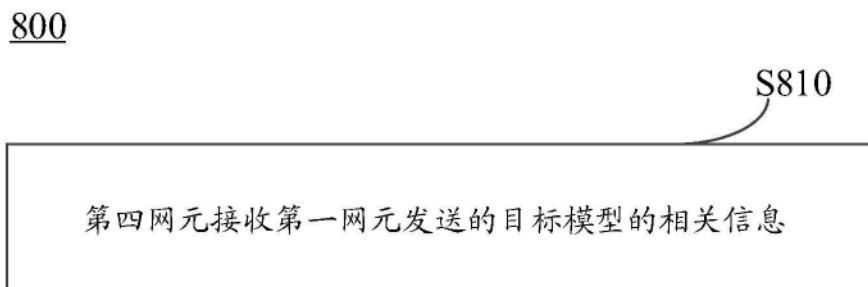


图8

900

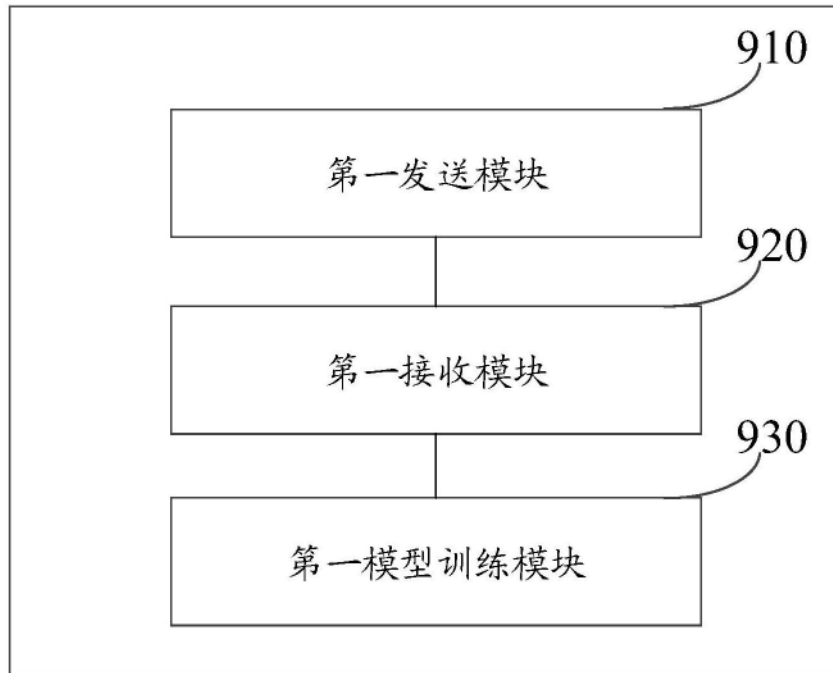


图9

1000

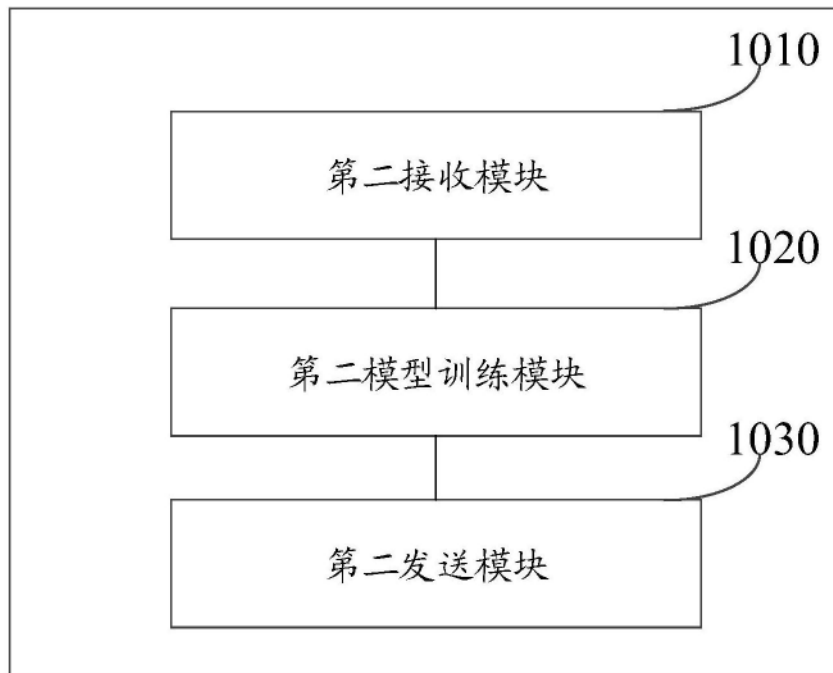


图10

1100

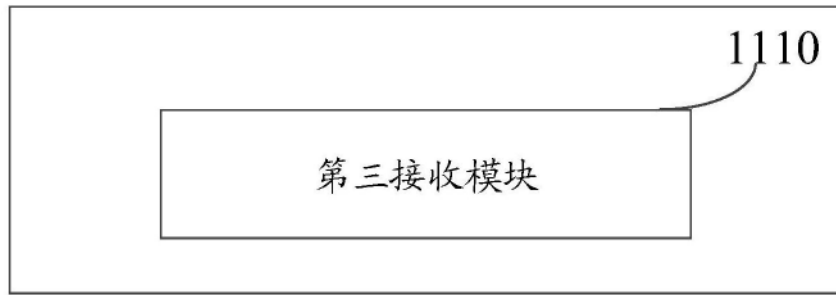


图11

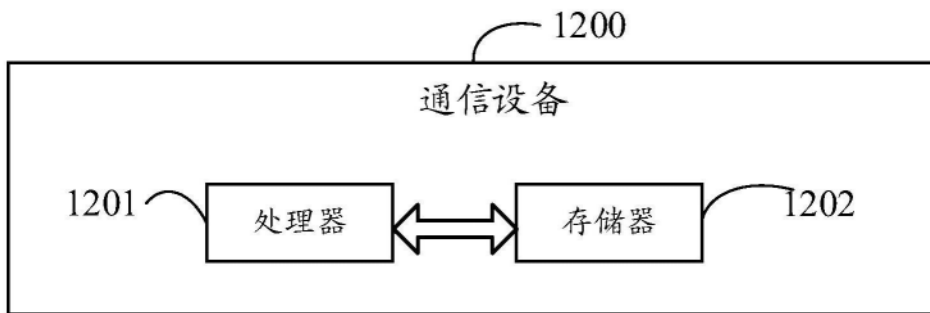


图12

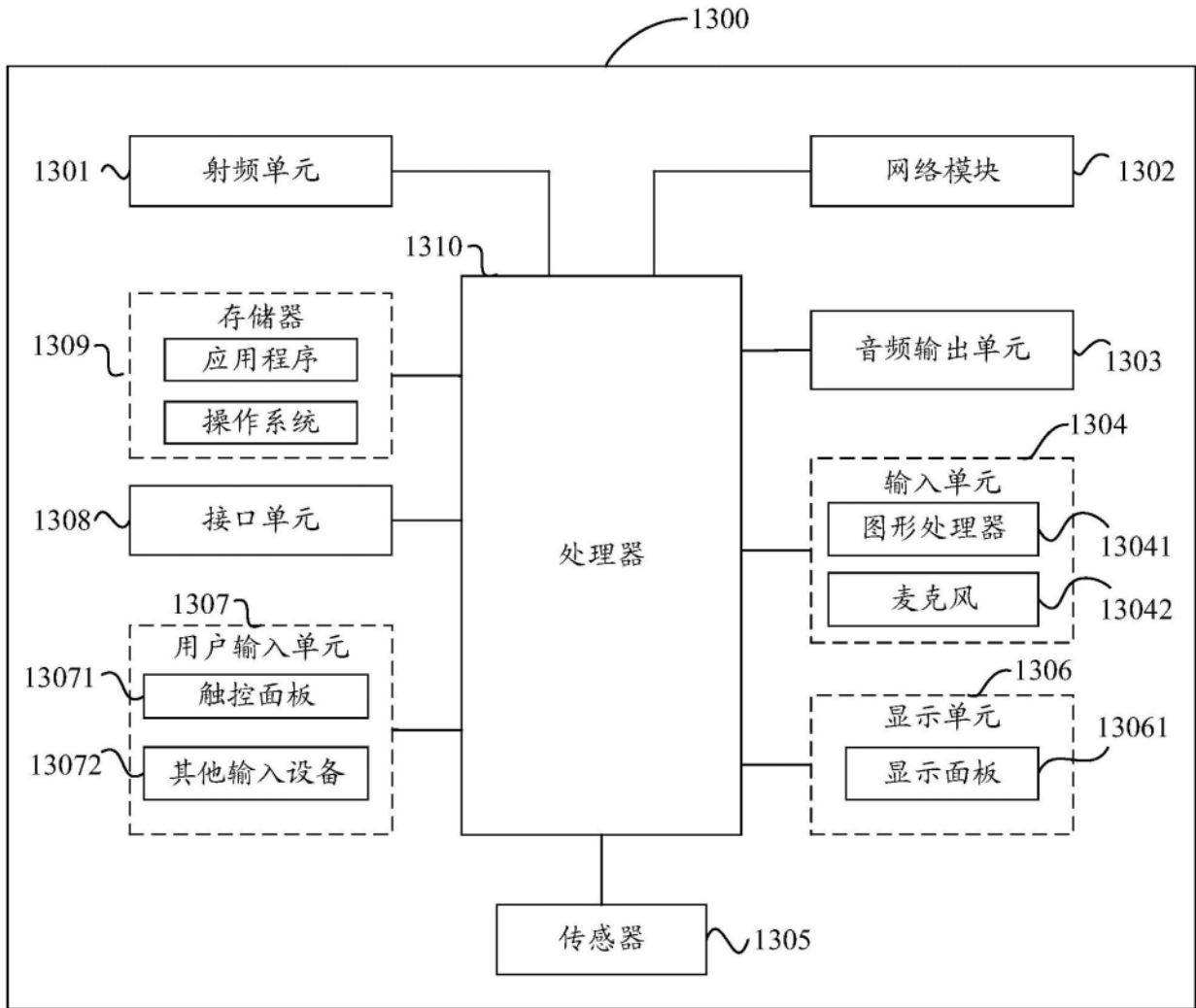


图13

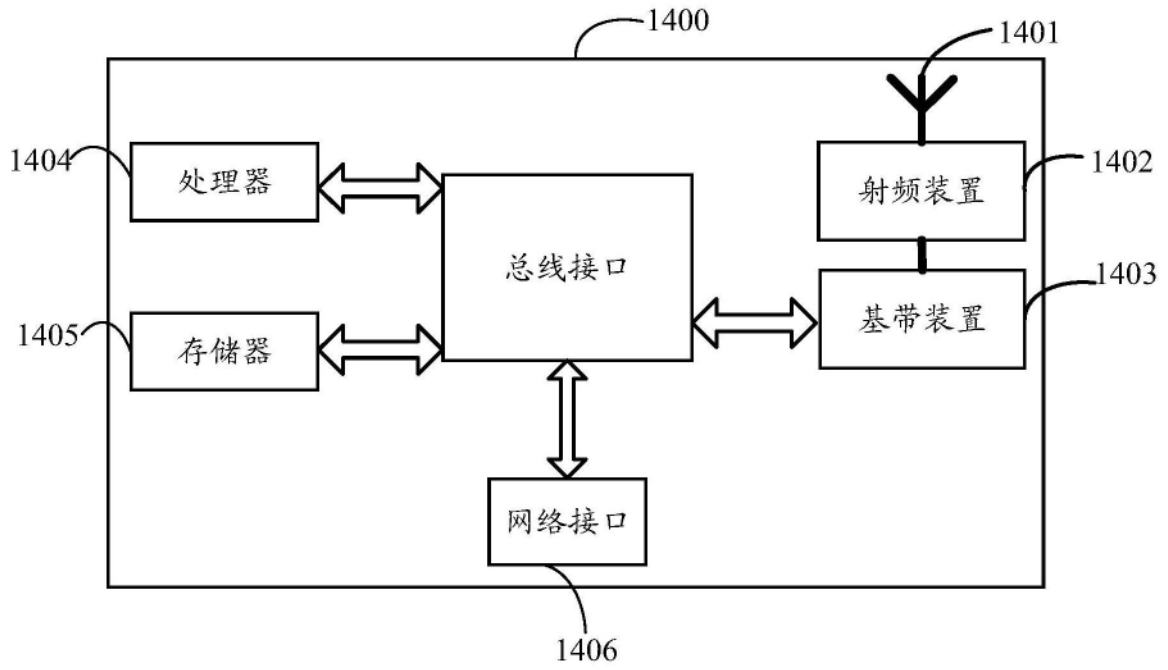


图14

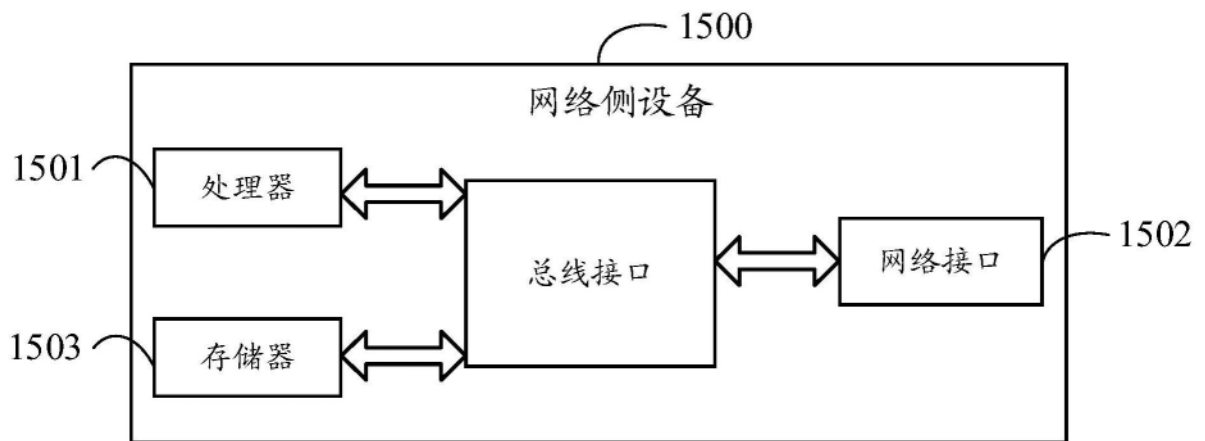


图15