



(21) 申请号 202180104709.4

(51) Int.Cl.

(22) 申请日 2021.12.15

H04W 4/50 (2006.01)

(85) PCT国际申请进入国家阶段日

2024.06.03

(86) PCT国际申请的申请数据

PCT/CN2021/138271 2021.12.15

(87) PCT国际申请的公布数据

W02023/108470 EN 2023.06.22

(71) 申请人 日本电气株式会社

地址 日本东京都

(72) 发明人 王达 梁林 王刚

(74) 专利代理机构 北京市金杜律师事务所

11256

专利代理师 李兴斌 程延霞

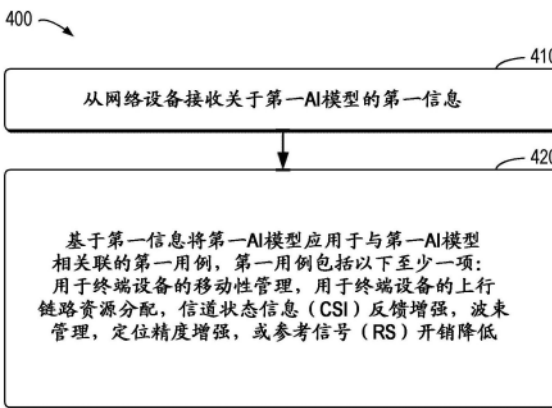
权利要求书5页 说明书11页 附图3页

(54) 发明名称

用于通信的方法、设备和计算机可读介质

(57) 摘要

本公开的实施例涉及用于通信的方法、设备和计算机可读介质。一种在终端设备处实现的方法,包括在终端设备处从网络设备接收关于第一人工智能(AI)模型的第一信息。该方法还包括基于第一信息将第一AI模型应用于与第一AI模型相关联的第一用例。第一用例包括以下至少一项:终端设备的移动性管理,终端设备的上行链路资源分配,信道状态信息(CSI)反馈增强,波束管理,定位精度增强,或参考信号(RS)开销降低。



1. 一种用于通信的方法,包括:
在终端设备处,从网络设备接收关于第一人工智能(AI)模型的第一信息;以及
基于所述第一信息,将所述第一AI模型应用于与所述第一AI模型相关联的第一用例,
所述第一用例包括以下至少一项:
所述终端设备的移动性管理,
所述终端设备的上行链路资源分配,
信道状态信息(CSI)反馈增强,
波束管理,
定位精度增强,或
参考信号(RS)开销降低。
2. 根据权利要求1所述的方法,进一步包括:
向所述网络设备发送关于至少一个AI模型的信息,
其中所述至少一个AI模型被预配置,并且包括所述第一AI模型,所述至少一个AI模型
中的每个AI模型与用于所述终端设备的至少一个用例相关联。
3. 根据权利要求2所述的方法,其中发送关于所述至少一个AI模型的所述信息包括:
将关于所述至少一个AI模型的所述信息与关于所述终端设备的能力信息一起进行发
送。
4. 根据权利要求1或2所述的方法,其中接收关于所述第一AI模型的所述第一信息包
括:
响应于所述第一用例将要被发起,接收关于所述第一AI模型的使能指示。
5. 根据权利要求4所述的方法,其中接收关于所述第一AI模型的所述使能指示包括:
经由无线电资源控制消息或系统信息,接收关于所述第一AI模型的所述使能指示。
6. 根据权利要求1所述的方法,其中应用所述第一AI模型包括:
将所述第一AI模型应用于所述移动性管理,以获得所述第一AI模型的第一输出,所述
第一输出与所述移动性管理相关联;并且
所述方法进一步包括:
向所述网络设备发送所述第一输出。
7. 根据权利要求6所述的方法,其中所述第一输出包括关于以下至少一项的信息:
至少一个被预测的候选小区,所述至少一个被预测的候选小区用于切换或主辅小区
(PSCe11)改变,
用于所述至少一个被预测的候选小区中的每个候选小区的被预测的执行条件,
用于所述切换或所述PSCe11改变的被预测的候选频率,
所述终端设备的被预测的轨迹,
所述终端设备的被预测的移动速度,或
所述终端设备的被预测的移动方向。
8. 根据权利要求1所述的方法,进一步包括:
从所述网络设备接收关于条件切换(CHO)或者条件主辅小区(PSCe11)添加或改变
(CPAC)的第二信息,所述第二信息指示用于所述CHO或所述CPAC的候选小区、并且指示所述
CHO或所述CPAC的执行条件与所述第一AI模型的第二输出相关联;并且

其中应用所述第一AI模型包括：

响应于基于所述第二输出来确定所述执行条件被满足，执行所述CHO或所述CPAC。

9. 根据权利要求8所述的方法，其中所述第二输出指示以下至少一项：

所述候选小区中的第一候选小区，或

所述终端设备对所述第一候选小区执行所述CHO或所述CPAC的概率。

10. 根据权利要求8所述的方法，其中应用所述第一AI模型包括：

响应于接收到所述第二信息，应用所述第一AI模型。

11. 根据权利要求1所述的方法，进一步包括：

从所述网络设备接收关于所述第一AI模型的有效性区域的第三信息；以及

应用所述第一AI模型包括：

在所述有效性区域内，将所述第一AI模型应用于空闲状态或非活动状态下的所述移动性管理。

12. 根据权利要求11所述的方法，其中所述有效性区域包括以下至少一项：

小区，

无线电接入网络通知区域，或

跟踪区。

13. 根据权利要求11所述的方法，进一步包括：

向所述网络设备发送关于所述第一AI模型的反馈信息，包括：

响应于从所述网络设备接收到对所述反馈信息的请求，发送所述反馈信息；或者

基于针对所述反馈信息的预配置来发送所述反馈信息。

14. 根据权利要求13所述的方法，其中所述反馈信息包括以下至少一项：

在所述第一AI模型被使能时关于所述终端设备的移动性历史信息，

关于用于空闲状态或非活动状态下的测量的功率的信息，

在所述第一AI模型被使能时与向小区的无线电资源控制 (RRC) 建立失败或RRC恢复失败相关的信息，或者

与所述第一AI模型被使能时所述终端设备完成向小区的RRC建立过程或RRC恢复过程之后不久执行切换的情况有关的信息。

15. 根据权利要求1所述的方法，进一步包括：

从所述网络设备接收关于所述第一AI模型的有效性定时器的第四信息；以及

应用所述第一AI模型包括：

在所述有效性定时器期满之前应用所述第一AI模型。

16. 根据权利要求1所述的方法，进一步包括：

从所述网络设备接收用于无线电承载或逻辑信道的服务质量 (QoS) 参数；

应用所述第一AI模型包括：

应用所述QoS参数作为所述第一AI模型的输入，以确定所述上行链路资源分配。

17. 根据权利要求16所述的方法，其中所述QoS参数包括用于所述无线电承载或逻辑信道的以下至少一项：

数据包延迟预算，

最大包错误率或丢包率，

保证比特率,
最大位速率,
优先比特率,
优先等级,
存活时间,或
第五代QoS标识符值。

18. 一种用于通信的方法,包括:

在网络设备处,确定关于第一人工智能(AI)模型的第一信息,所述第一AI模型与第一用例相关联,所述第一用例包括以下至少一项:

终端设备的移动性管理,
所述终端设备的上行链路资源分配,
信道状态信息(CSI)反馈增强,
波束管理,
定位精度增强,或
参考信号(RS)开销降低;

向所述终端设备发送关于所述至少一个AI模型的所述第一信息。

19. 根据权利要求18所述的方法,进一步包括:

从所述终端设备接收关于至少一个AI模型的信息;以及
其中所述至少一个AI模型被预配置,并且包括所述第一AI模型,所述至少一个AI模型中的每个AI模型与用于所述终端设备的至少一个用例相关联。

20. 根据权利要求19所述的方法,其中接收关于所述至少一个AI模型的所述信息包括:
将关于所述至少一个AI模型的所述信息与关于所述终端设备的能力信息一起进行接收。

21. 根据权利要求18或19所述的方法,其中发送关于所述第一AI模型的所述第一信息包括:

响应于所述第一用例将要被发起,发送关于所述第一AI模型的使能指示。

22. 根据权利要求21所述的方法,其中发送关于所述第一AI模型的所述使能指示包括:
经由无线电资源控制消息或系统信息,发送关于所述第一AI模型的所述使能指示。

23. 根据权利要求18所述的方法,其中所述第一AI模型被应用于所述移动性管理,以获得所述第一AI模型的第一输出,所述第一输出与所述移动性管理相关联;以及

所述方法进一步包括:

从所述终端设备接收所述第一输出。

24. 根据权利要求23所述的方法,其中所述第一输出包括关于以下至少一项的信息:
至少一个被预测的候选小区,所述至少一个被预测的候选小区用于切换或主辅小区(PSCell)改变,

用于所述至少一个被预测的候选小区中的每个候选小区的被预测的执行条件,

用于所述切换或所述PSCell改变的被预测的候选频率,

所述终端设备的被预测的轨迹,

所述终端设备的被预测的移动速度,或

所述终端设备的被预测的移动方向。

25. 根据权利要求18所述的方法,进一步包括:

向所述终端设备发送关于条件切换 (CHO) 或者条件主辅小区 (PSCe11) 添加或改变 (CPAC) 的第二信息,所述第二信息指示用于所述CHO或所述CPAC的候选小区,并且指示所述CHO或所述CPAC的执行条件与所述第一AI模型的第二输出相关联。

26. 根据权利要求25所述的方法,其中所述第二输出指示以下至少一项:

所述候选小区中的第一候选小区,或

所述终端设备对所述第一候选小区执行所述CHO或所述CPAC的概率。

27. 根据权利要求18所述的方法,进一步包括:

向所述终端设备发送关于所述第一AI模型的有效性区域的第三信息。

28. 根据权利要求27所述的方法,其中所述有效性区域包括以下至少一项:

小区,

无线电接入网络通知区域,或

跟踪区。

29. 根据权利要求27所述的方法,进一步包括:

从所述终端设备接收关于所述第一AI模型的反馈信息,包括:

向所述终端设备发送对所述反馈信息的请求;以及

基于所述请求来接收所述反馈信息。

30. 根据权利要求27所述的方法,进一步包括:

从所述终端设备接收关于所述第一AI模型的反馈信息,包括:

基于针对所述反馈信息的预配置来接收所述反馈信息。

31. 根据权利要求29或30所述的方法,其中所述反馈信息包括以下至少一项:

在所述第一AI模型被使能时关于所述终端设备的移动性历史信息,

关于用于空闲状态或非活动状态下的测量的功率的信息,

在所述第一AI模型被使能时与向小区的无线电资源控制 (RRC) 建立失败或RRC恢复失败相关的信息,或

与所述第一AI模型被使能时所述终端设备完成向小区的RRC建立过程或RRC恢复过程之后不久执行切换的情况有关的信息。

32. 根据权利要求18所述的方法,进一步包括:

向所述终端设备发送关于所述第一AI模型的有效性定时器的第四信息。

33. 根据权利要求18所述的方法,进一步包括:

向所述终端设备发送用于无线承载或逻辑信道的服务质量 (QoS) 参数。

34. 根据权利要求33所述的方法,其中所述QoS参数包括用于所述无线电承载或逻辑信道的以下至少一项:

数据包延迟预算,

最大包错误率或丢包率,

保证比特率,

最大位速率,

优先比特率,

优先等级,
存活时间,或
第五代QoS标识符值。

35.一种终端设备,包括:

处理器;以及

存储器,所述存储器耦合到所述处理器并且其上存储有指令,所述指令在由所述处理器执行时,使所述终端设备执行根据权利要求1至17中任一项所述的方法。

36.一种网络设备,包括:

处理器;以及

存储器,所述存储器耦合到所述处理器并且其上存储有指令,所述指令在由所述处理器执行时,使所述网络设备执行根据权利要求18至34中任一项所述的方法。

37.一种计算机可读介质,其上存储有指令,所述指令当在至少一个处理器上被执行时,使所述至少一个处理器执行根据权利要求1至17中的任一项、或者权利要求18至34中的任一项所述的方法。

用于通信的方法、设备和计算机可读介质

技术领域

[0001] 本公开的实施例一般涉及电信领域,尤其涉及用于通信的方法、设备和计算机可读介质。

背景技术

[0002] 第五代(5G)网络被预期满足增加数目的关键性能指标(KPI)的一致优化的挑战,该KPI包括时延、可靠性、连接密度、用户体验、能量效率等。人工智能(AI)或机器学习(ML)提供了强大的工具来帮助运营商通过分析所收集的和自主处理的数据来改进网络管理和用户体验,这可以产生进一步的洞察。

[0003] 第三代合作伙伴计划(3GPP)现在正在空中接口上努力改进,其特征使得能够改进对基于AI/ML的算法的支持,以提高性能和/或降低复杂度或开销。增强的性能可以取决于所考虑的用例,并且可以是改进的吞吐量、鲁棒性、准确性或可靠性,或减少的开销等。

发明内容

[0004] 一般而言,本公开的示例实施例提供了用于通信的方法、设备和计算机可读介质。

[0005] 在第一方面,提供了一种用于在终端设备处实现的通信的方法。该方法包括在终端设备处从网络设备接收关于第一人工智能(AI)模型的第一信息。该方法还包括基于第一信息将第一AI模型应用于与第一AI模型相关联的第一用例。第一用例包括以下至少一项:终端设备的移动性管理,终端设备的上行链路资源分配,信道状态信息(CSI)反馈增强,波束管理,定位精度增强,或参考信号(RS)开销降低。

[0006] 在第二方面,提供了一种用于在网络设备处实现的通信的方法。该方法包括在网络设备处确定关于与第一用例相关联的第一人工智能(AI)模型的第一信息。第一用例包括以下至少一项:终端设备的移动性管理,终端设备的上行链路资源分配,信道状态信息(CSI)反馈增强,波束管理,定位精度增强,或参考信号(RS)开销降低。该方法还包括将关于第一AI模型的第一信息发送到终端设备。

[0007] 在第三方面,提供了一种终端设备。终端设备包括处理器和存储指令的存储器。存储器和指令被配置为与处理器一起使终端设备执行根据第一方面的方法。

[0008] 在第四方面,提供了一种网络设备。网络设备包括处理器和存储指令的存储器。存储器和指令被配置为与处理器一起使网络设备执行根据第二方面的方法。

[0009] 在第五方面,提供了一种其上存储有指令的计算机可读介质。指令当在设备的至少一个处理器上执行时使设备执行根据第一方面的方法。

[0010] 在第六方面,提供了一种其上存储有指令的计算机可读介质。指令当在设备的至少一个处理器上执行时使设备执行根据第二方面的方法。

[0011] 应当理解,概述部分不旨在标识本公开的实施例的关键或必要特征,也不旨在用于限制本公开的范围。通过以下描述,本公开的其他特征将变得容易理解。

附图说明

[0012] 通过在附图中的本公开的一些实施例的更详细的描述,本公开的上述和其他目的,特征和优点将变得更加显而易见,其中:

[0013] 图1图示了可以实现本公开的实现方式的示例通信网络;

[0014] 图2图示了根据本公开的一些实施例的AI模型的示例;

[0015] 图3图示了示出根据本公开的一些实施例的用于在终端设备处使用AI模型的示例过程的示例信令图;

[0016] 图4图示了根据本公开的一些实施例的示例方法的流程图;

[0017] 图5图示了根据本公开的一些其他实施例的示例方法的流程图;以及

[0018] 图6是适于实现本公开的实施例的设备的简化框图。

[0019] 在所有附图中,相同或相似的附图标记表示相同或相似的元件。

具体实施方式

[0020] 现在将参考一些示例实施例来描述本公开的原理。应当理解,描述这些实施例仅是为了说明的目的,并且帮助本领域技术人员理解和实现本公开,而不暗示对公开的范围的任何限制。这里描述的公开可以以不同于下面描述的方式的各种方式来实现。

[0021] 在以下描述和权利要求中,除非另外定义,否则本文使用的所有技术和科学术语具有与本公开所属领域的普通技术人员通常理解相同含义。

[0022] 如这里所使用的,术语“终端设备”是指具有无线或有线通信能力的任何设备。终端设备的示例包括但不限于用户设备(UE),个人计算机,桌面,移动电话,蜂窝电话,智能电话,个人数字助理(PDA),

便携式计算机,平板,可穿戴设备,物联网(IoT)设备,超可靠和低等待时间通信(URLLC)设备,各种物联网(IoE)设备,机器类型通信(MTC)设备,用于V2X通信的车载设备(其中X表示行人,车辆或基础设施/网络),用于集成接入和回传(IAB)的设备,包括卫星和包括无人飞行器系统(UAS)的高空平台(HAP)的非陆地网络(NTN)中的空运交通工具或空中交通工具,包括诸如增强现实(AR)、混合现实(MR)和虚拟现实(VR)的不同类型的现实的扩展现实(XR)设备,通常被称为无人驾驶飞机的无人驾驶飞行器(UAV)(该无人驾驶飞机是没有任何人工飞行员的飞机),高速列车(HST)上的设备,或诸如数码相机,传感器,游戏设备,音乐存储和回放设备的图像捕获设备,或允许无线或有线因特网访问和浏览的因特网设备等。“终端设备”可以进一步具有“多播/广播”特征,以支持公共安全和关键任务,V2X应用,透明IPv4/IPv6多播递送,IPTV,智能TV,无线电服务,无线上的软件递送,群组通信和IoT应用。它还可以结合一个或多个用户识别模块(SIM),如公知的多SIM。术语“终端设备”可以与UE,移动站,订户站,移动终端,用户终端或无线设备互换使用。

[0023] 如这里所使用的,术语“网络设备”指的是能够提供或主控终端设备可以通信的小区或覆盖的设备。网络设备的示例包括但不限于节点B(节点B或NB),演进型节点B(eNodeB或eNB),下一代节点B(gNB),传输接收点(TRP),远程无线电单元(RRU),无线电头(RH),远程无线电头(RRH),IAB节点,诸如毫微微节点的低功率节点,微微节点,可重新配置智能表面(RIS)等。

[0024] 终端设备或网络设备可以具有人工智能(AI)或机器学习能力。它通常包括根据为

特定功能收集的大量数据训练的模型,并可用于预测一些信息。

[0025] 终端或网络设备可以在几个频率范围上工作,例如FR1 (410MHz—7125MHz),FR2 (24.25GHz到71GHz),大于100GHz的频带以及太赫兹 (THz)。它还可以在许可/未许可/共享频谱上工作。在多无线电双连接 (MR—DC) 应用场景下,终端设备可以具有与网络设备的多于一个连接。终端设备或网络设备可以工作在全双工,灵活双工和交叉划分双工模式。

[0026] 本公开的实施例可以在测试设备中执行,例如信号发生器,信号分析器,频谱分析器,网络分析器,测试终端设备,测试网络设备,信道仿真器。

[0027] 如本文所用,单数形式“一”,“一个”和“该”也旨在包括复数形式,除非上下文另外明确指出。术语“包括”及其变体将被解读为开放式术语,其意味着“包括但不限于”。术语“基于”将被解读为“至少部分地基于”。术语“一些实施例”和“实施例”将被解读为“至少一些实施例”。术语“另一实施例”将被解读为“至少一个其他实施例”。

[0028] 在一些示例中,值,过程或装置被称为“最佳”,“最低”,“最高”,“最小”,“最大”等。应当理解,这样的描述旨在表示可以在许多所使用的功能备选方案中进行选择,并且这样的选择不需要比其它选择更好,更小,更高或更优选。

[0029] 图1图示了其中可以实现本公开的实施例的示例通信网络100。网络100包括终端设备110和服务于终端设备110的网络设备120。网络设备120的服务区域被称为小区102。应当理解,网络设备和终端设备的数目仅用于说明的目的,而不暗示任何限制。系统100可以包括适合于实现本公开的实施例的任何合适数目的网络设备和终端设备。尽管未示出,但是应当理解,一个或多个终端设备可以位于小区102中并由网络设备120服务。

[0030] 通信网络100中的通信可根据当前已知或将来开发的任何一代通信协议来实现。通信协议的示例包括但不限于第一代 (1G)、第二代 (2G)、2.5G、2.75G、第三代 (3G)、第四代 (4G)、4.5G、第五代 (5G) 通信协议,5.5G、5G高级网络或第六代 (6G) 网络。

[0031] 终端设备110和网络设备120中的至少一个可以具有AI或ML能力。通常,从针对特定功能的大量收集的数据训练的AI模型可用于预测一些信息。

[0032] 图2图示了根据本公开的一些实施例的AI模型200的示例。如图所示,AI模型200包括数据收集功能210,模型训练功能220,模型推理功能230和行动者 (actor) 功能240。

[0033] 数据收集功能210可以是向模型训练功能220和模型推理功能230提供输入数据的功能。输入数据的示例可以包括来自终端设备或不同网络实体的测量、来自行动者功能240的反馈、来自AI模型200的输出。

[0034] 模型训练功能220可以是执行AI模型200的训练、验证和测试的功能。如果需要,那么模型训练功能220还可以负责基于由数据收集功能210递送的训练数据进行数据准备 (例如,数据预处理和清洗,格式化和变换)。

[0035] 模型推理功能230可以是提供AI模型200的推理输出 (例如预测或决策) 的功能。在下文中,为了简洁起见,“推理输出”也将被称为“输出”。如果需要,那么模型推理功能230还可以负责基于由数据收集功能210递送的推理数据进行数据准备 (例如,数据预处理和清洗,格式化和变换)。

[0036] 行动者功能240可以是接收来自模型推理功能230的输出并触发或执行相应动作的功能。行动者功能240可以触发指向其他实体或其自身的动作。

[0037] 行动者功能240还可以向数据收集功能210提供反馈。反馈可以包括导出训练或推

理数据或性能反馈所需的信息。

[0038] 在本公开中,模型推理功能230由终端设备110执行。在一些实施例中,模型训练功能220可以由网络设备120执行。在这样的实施例中,网络设备120可以将模型推理功能230配置给终端设备110。在一些其它实施例中,模型训练功能220可由图1中未示出的另一网络设备执行。例如,模型训练功能220可以由运营和维护(OAM)设备来执行。在这样的实施例中,另一网络设备可以将模型推理功能230配置给终端设备110。

[0039] 在其他实施例中,可以在终端设备110处预先配置至少一个AI模型。至少一个AI模型中的每个AI模型可以与终端设备110的至少一个用例相关联。终端设备110可以将关于至少一个AI模型的信息发送到网络设备120。例如,关于至少一个AI模型的信息可以指示终端设备110是否支持至少一个AI模型。又例如,关于至少一个AI模型的信息可以指示终端设备110支持的至少一个AI模型以及与至少一个AI模型中的每一个AI模型相关联的至少一个用例。

[0040] 在一些实施例中,响应于用例要被发起,网络设备120可以向终端设备110发送关于AI模型的使能指示。

[0041] 传统上,终端设备的移动性管理基于网络设备处的AI模型,这导致无法收集用于移动性决策的即时信息。此外,终端设备的上行链路资源分配基于令牌桶机制。然而,该机制仅考虑有限的因素,这使得输出不是最优的。

[0042] 本公开的实施例提供了一种用于在终端设备处使用AI模型来解决上述问题和一个或多个其他潜在问题的方案。根据该方案,终端设备从网络设备接收关于第一AI模型的第一信息,并基于第一信息将第一AI模型应用于与第一AI模型相关联的第一用例。第一用例包括以下至少一项:终端设备的移动性管理,终端设备的上行链路资源分配,信道状态信息(CSI)反馈增强,波束管理,定位精度增强,或参考信号(RS)开销降低。以此方式,可以收集用于移动性决策的即时信息。此外,可以实现基于AI的上行链路资源分配。

[0043] 现在将参照图3至图5描述本发明的原理。现在参考图3,图3图示了示出根据本公开的一些示例实施例的用于在终端设备处使用AI模型的过程300的信令图。为了讨论的目的,将参考图1描述过程300。过程300可以涉及如图1所示的终端设备110和网络设备120。尽管已经在图1的通信网络100中描述了过程300,但是该过程同样可以应用于其它通信场景。

[0044] 如图3所示,终端设备110从网络设备120接收(320)关于第一AI模型的第一信息。

[0045] 接着,终端设备110基于第一信息将第一AI模型应用(330)于与第一AI模型相关联的第一用例。第一用例包括以下至少一项:终端设备110的移动性管理,终端设备110的上行链路资源分配,信道状态信息(CSI)反馈增强,波束管理,定位精度增强或参考信号(RS)开销降低。

[0046] 根据本公开,因为第一AI模型位于终端设备110处,所以可以收集用于移动性决策的即时信息。此外,可以实现基于AI的上行链路资源分配。

[0047] 在一些实施例中,至少一个AI模型可以在终端设备110处被预先配置,并且可以包括第一AI模型。至少一个AI模型中的每个AI模型与终端设备110的至少一个用例相关联。在这样的实施例中,终端设备110可以向网络设备120发送(310)关于至少一个AI模型的信息。例如,终端设备110可以将关于至少一个AI模型的信息与关于终端设备110的能力信息一起进行发送。

[0048] 在至少一个AI模型在终端设备110处被预先配置的实施例中,关于第一AI模型的第一信息可以包括关于第一AI模型的使能指示。在这样的实施例中,终端设备110可以响应于第一用例要被发起而接收关于第一AI模型的使能指示。例如,终端设备110可以经由无线电资源控制(RRC)消息或系统信息来接收使能指示。

[0049] 在至少一个AI模型没有在终端设备110处被预先配置的实施例中,网络设备120可以将第一AI模型配置给终端设备110。在这样的实施例中,关于第一AI模型的第一信息可以包括关于第一AI模型的配置信息。例如,关于第一AI模型的第一信息可以包括关于第一AI模型的模型推理功能的配置信息。在这样的实施例中,网络设备120可以在安全性已经被激活之后使用RRC消息来配置第一AI模型。

[0050] 在网络设备120配置第一AI模型的实施例中,第一AI模型可以作为一个容器被递送到终端设备110,该容器的内容对于RRC层是透明的。在网络设备120充当主节点(MN)的实施例中,MN可以通过信令无线电承载(SRB)(例如SRB1或SRB2)向终端设备110递送第一AI模型。在网络设备120充当辅节点(SN)的实施例中,SN可以通过SRB(例如SRB3)向终端设备110递送第一AI模型。可替换地,SN可以向MN发送第一AI模型,并且MN使用SRB(例如SRB1或SRB2)向终端设备110递送第一AI模型。可替换地,可以使用一个或多个专用于AI模型配置的新SRB。

[0051] 在一些实施例中,如果终端设备110过热或存储器不足,则终端设备110可以通过RRC消息请求网络设备120释放一个或多个AI模型。终端设备110可以使用诸如UE Assistance Information之类的RRC消息来请求释放一个或多个AI模型,并且诸如过热或存储器不足这样的释放原因可以被包括在该消息中。

[0052] 在一些实施例中,终端设备110可以在接收到第一AI模型的配置信息或使能指示后应用第一AI模型。

[0053] 在一些实施例中,终端设备110可以直接应用第一AI模型的输出。可替换地,终端设备110可以向配置第一AI模型的网络设备报告第一AI模型的输出。应当注意,配置第一AI模型的网络设备可以与网络设备120相同或不同。这样,网络设备可以根据来自终端设备110的输出来动作,或者使用来自终端设备110的输出作为网络设备处的AI模型(推断或训练)的输入。

[0054] 在一些实施例中,终端设备110可以在容器中向网络设备报告第一AI模型的输出,该容器的内容对于RRC层是透明的。

[0055] 在一些实施例中,终端设备110可以通过指定的RRC IE或专用消息向网络设备发送(340)关于第一AI模型的反馈信息。

[0056] 在一些实施例中,如果第一AI模型的模型训练功能位于网络侧,例如位于网络设备120或OAM侧,则网络还可以配置或请求终端设备110向网络发送关于第一AI模型的反馈信息。在这样的实施例中,网络可以使用反馈信息作为模型训练功能的输入。

[0057] 当前,对于处于连接状态的终端设备,移动性管理(例如,切换,主辅小区(PSCell)改变)取决于网络决策。在传统中,这取决于gNB实现。目前,在版本(Release)17RAN3 AI/ML研究项目中,正在研究用于移动性性能优化的基于网络的AI/ML。然而,网络AI/ML模型基于来自终端设备的反馈信息,可能存在延迟并且大量的信息可以被报告给网络。

[0058] 为了解决上述问题,在一些实施例中,终端设备110可以将第一AI模型应用于移动

性管理,以便获得第一AI模型的第一输出。第一输出与移动性管理相关联。终端设备110进而可以将第一输出发送到网络设备120。网络设备120可以在最终的移动性配置决策中考虑第一输出。例如,终端设备110可以通过诸如UE Assistance Information的RRC消息发送第一输出。这样,可以收集用于移动性决策的即时信息,并且与传统的移动性过程相比,该即时信息更加准确和有效。

[0059] 在一些实施例中,该第一输出可以包括关于以下至少一项的信息:用于切换或PSCell改变的至少一个被预测的候选小区,用于至少一个被预测的候选小区中的每个候选小区的被预测的执行条件,用于切换或PSCell改变的被预测的候选频率,终端设备110的被预测的轨迹,终端设备110的被预测的移动速度,或终端设备110的被预测的移动方向。

[0060] 条件切换(CHO)被定义为当满足一个或多个切换执行条件时由终端设备执行的切换。终端设备在接收到CHO配置时开始评估一个或多个执行条件,并且一旦执行切换(诸如传统切换或条件切换执行)就停止评估一个或多个执行条件。类似地,条件PSCell添加/改变(CPAC)被定义为当满足至少一个执行条件时执行的PSCell添加或改变。

[0061] 在一些实施例中,终端设备110可以从网络设备120接收关于CHO或CPAC的第二信息。第二信息指示CHO或CPAC的候选小区,并指示CHO或CPAC的执行条件与第一AI模型的第二输出相关联。在这样的实施例中,终端设备110可以响应于基于第二输出确定执行条件被满足而执行CHO或CPAC。

[0062] 在这样的实施例中,第二输出可以指示以下至少一项:候选小区中的第一候选小区,或者终端设备110对第一候选小区执行CHO或CPAC的概率。

[0063] 在这样的实施例中,终端设备110可以响应于接收到第二信息而应用第一AI模型。

[0064] 在这样的实施例中,为了节省功耗,在执行CHO或CPAC时,终端设备110可以停止第一AI模型。

[0065] 当前,空闲状态或非活动状态下的终端设备的移动性管理基于小区重选。小区重选基于许多因素,例如,空闲或非活动测量,频率优先级,服务,切片。终端设备的当前行为可能导致驻留在不是非常适合的小区上,并且网络必须在终端设备接入该小区之后几乎立即将终端设备切换到另一小区。

[0066] 为了解决上述问题,在一些实施例中,第一AI模型可以是用于小区重选的AI模型或用于空闲/非活动状态测量放松的AI模型。在这样的实施例中,终端设备110可以从网络设备120接收关于第一AI模型的有效性区域的第三信息。终端设备110进而可以在有效性区域内将第一AI模型应用于空闲或非活动状态下的移动性管理。

[0067] 在一些实施例中,如果终端设备110重选到不属于有效性区域的小区,则终端设备110可以不应用第一模型,暂停第一模型或释放第一模型。

[0068] 在一些实施例中,有效性区域包括以下至少一项:小区,无线电接入网通知(RNA)区域或跟踪区域。

[0069] 在这样的实施例中,响应于从网络设备120接收到对反馈信息的请求,终端设备110可以向网络设备120发送关于第一AI模型的反馈信息。可替换地,终端设备110可以基于反馈信息的预配置来发送关于第一AI模型的反馈信息。

[0070] 在这样的实施例中,反馈信息可以包括以下至少一项:在第一AI模型被使能时关于终端设备110的移动性历史信息,关于用于空闲状态或非活动状态下的测量的功率的信

息,在第一AI模型被使能时与向小区的RRC建立失败或RRC恢复失败相关的信息,或者与第一AI模型被使能时在终端设备110完成向小区的RRC建立过程或RRC恢复过程之后不久执行切换的情况有关的信息。

[0071] 例如,关于终端设备110的移动性历史信息可以包括关于终端设备110的轨迹和所驻留的小区的信息。关于用于空闲状态或非活动状态下的测量的功率的信息可以包括终端设备110的功率使用水平。与向小区的RRC建立失败或RRC恢复失败相关的信息可以指示在连接建立失败报告(也称为ConnEstFailReport)中添加AI信息。

[0072] 在一些实施例中,终端设备110可以从网络设备120接收关于第一AI模型的有效性定时器的第四信息。进而,终端设备110可以在有效性定时器期满之前应用第一AI模型。在这样的实施例中,在接收到第一AI模型的配置或接收到第一AI模型的使能指示时有效性定时器启动。如果有效性定时器期满,则终端设备110不应用、暂停或释放第一AI模型。

[0073] 当前,上行链路资源分配基于令牌桶机制。然而,这种机制只考虑有限的因素,这使得输出不是最优的。

[0074] 为了解决上述问题,在一些实施例中,终端设备110可以从网络设备120接收用于无线电承载或逻辑信道的服务质量(QoS)参数。接着,终端设备110可以应用QoS参数作为第一AI模型的输入,以确定上行链路资源分配。这样,可以实现基于AI的上行链路资源分配。

[0075] 在一些实施例中,QoS参数可以包括用于无线承载或逻辑信道的以下各项中的至少一项:数据包延迟预算,最大包错误率或丢包率,保证比特率,最大位速率,优先比特率,优先等级,存活时间,或第五代QoS标识符(5QI)值。

[0076] 图4图示了根据本公开的一些实施例的示例方法400的流程图。在一些实施例中,方法400可以在终端设备处实现。例如,方法400可以在如图1所示的终端设备110处实现。

[0077] 在框410,终端设备110从网络设备接收关于第一AI模型的第一信息。

[0078] 在框420,终端设备110基于第一信息将第一AI模型应用于与第一AI模型相关联的第一用例。第一用例包括以下至少一项:用于终端设备的移动性管理,用于终端设备的上行链路资源分配,信道状态信息(CSI)反馈增强,波束管理,定位精度增强,或参考信号(RS)开销降低。

[0079] 在一些实施例中,至少一个AI模型可以是预配置的并且包括第一AI模型,至少一个AI模型中的每一个可以与终端设备的至少一个用例相关联。终端设备110将关于至少一个AI模型的信息发送到网络设备。

[0080] 在一些实施例中,终端设备110可以将关于至少一个AI模型的信息与关于终端设备110的能力信息一起发送。

[0081] 在一些实施例中,响应于第一用例将要被发起,终端设备110可以接收关于第一AI模型的使能指示。

[0082] 在一些实施例中,终端设备110可以经由无线电资源控制消息或系统信息来接收关于第一AI模型的使能指示。

[0083] 在一些实施例中,终端设备110可以将第一AI模型应用于移动性管理,以获得第一AI模型的第一输出,该第一输出与移动性管理相关联。终端设备110进而可以将第一输出发送到网络设备。

[0084] 在一些实施例中,该第一输出包括关于以下各项中的至少一项的信息:至少一个

被预测的候选小区,至少一个被预测的候选小区用于切换或主辅小区 (PSCell) 改变,用于至少一个被预测的候选小区中的每个候选小区的被预测的执行条件,用于切换或PSCell改变的被预测的候选频率,终端设备的被预测的轨迹,终端设备的被预测的移动速度,或终端设备的被预测的移动方向。

[0085] 在一些实施例中,终端设备110可以从网络设备接收关于条件切换 (CHO) 或条件主辅小区 (PSCell) 添加或改变 (CPAC) 的第二信息。第二信息指示用于CHO或CPAC的候选小区,并指示CHO或CPAC的执行条件与第一AI模型的第二输出相关联。在这样的实施例中,终端设备110可以响应于基于第二输出确定执行条件被满足而执行CHO或CPAC。

[0086] 在一些实施例中,第二输出指示以下至少一项:候选小区中的第一候选小区,或者终端设备对第一候选小区执行CHO或CPAC的概率。

[0087] 在一些实施例中,终端设备110可以响应于接收到第二信息而应用第一AI模型。

[0088] 在一些实施例中,终端设备110可以从网络设备接收关于第一AI模型的有效性区域的第三信息,并且终端设备110可以在有效性区域内将第一AI模型应用于空闲或非活动状态下的移动性管理。

[0089] 在一些实施例中,有效性区域包括以下至少一项:小区,无线电接入网通知区域或跟踪区域。

[0090] 在一些实施例中,响应于从网络设备接收到对反馈信息的请求,终端设备110可以向网络设备发送关于第一AI模型的反馈信息。可替换地,终端设备110可以基于反馈信息的预配置来发送反馈信息。

[0091] 在一些实施例中,该反馈信息包括以下至少一项:在第一AI模型被使能时关于终端设备的移动性历史信息,关于用于空闲状态或非活动状态下的测量的功率的信息,在第一AI模型被使能时与向小区的无线电资源控制 (RRC) 建立失败或RRC恢复失败相关的信息,或与第一AI模型被使能时在终端设备完成向小区的RRC建立过程或RRC恢复过程之后不久执行切换的情况有关的信息。

[0092] 在一些实施例中,终端设备110可以从网络设备接收关于第一AI模型的有效性定时器的第四信息,并且终端设备110可以在有效性定时器期满之前应用第一AI模型。

[0093] 在一些实施例中,终端设备110可以从网络设备接收用于无线电承载或逻辑信道的服务质量 (QoS) 参数,并且终端设备110可以应用QoS参数作为第一AI模型的输入来确定上行链路资源分配。

[0094] 在一些实施例中,QoS参数包括用于无线电承载或逻辑信道的以下至少一项:数据包延迟预算,最大包错误率或丢包率,保证比特率,最大位速率,优先比特率,优先等级,存活时间,或第五代QoS标识符值。

[0095] 图5图示了根据本公开的一些实施例的示例方法500的流程图。在一些实施例中,方法500可以在网络设备处实现。例如,方法500可以在如图1所示的网络设备120处实现。

[0096] 在框510,网络设备120确定关于与第一用例相关联的第一AI模型的第一信息。第一用例包括以下至少一项:用于终端设备的移动性管理,用于终端设备的上行链路资源分配,信道状态信息 (CSI) 反馈增强,波束管理,定位精度增强,或参考信号 (RS) 开销降低。

[0097] 在框520,网络设备120将关于第一AI模型的第一信息发送到终端设备。

[0098] 在一些实施方式中,至少一个AI模型是预配置的并且包括第一AI模型。至少一个

AI模型中的每一个与终端设备的至少一个用例相关联。在这样的实施例中,网络设备120可以从终端设备接收关于至少一个AI模型的信息。

[0099] 在一些实施例中,网络设备120可以将关于至少一个AI模型的信息与关于终端设备的能力信息一起进行接收。

[0100] 在一些实施例中,响应于第一用例将要被发起,网络设备120可以向终端设备发送关于第一AI模型的使能指示。

[0101] 在一些实施例中,网络设备120可以经由无线电资源控制消息或系统信息来发送关于第一AI模型的使能指示。

[0102] 在一些实施例中,第一AI模型被应用于移动性管理,以获得第一AI模型的第一输出,该第一输出与移动性管理相关联。在这样的实施例中,网络设备120可以从终端设备接收第一输出。

[0103] 在一些实施例中,该第一输出包括关于以下各项中的至少一项的信息:用于切换或主辅小区(PSCe11)改变的至少一个被预测的候选小区,用于至少一个被预测的候选小区中的每个候选小区的被预测的执行条件,用于切换或PSCe11改变的被预测的候选频率,终端设备的被预测的轨迹,终端设备的被预测的移动速度,或终端设备的被预测的移动方向。

[0104] 在一些实施例中,网络设备120可以向终端设备发送关于条件切换(CHO)或者条件主辅小区(PSCe11)添加或改变(CPAC)的第二信息。第二信息指示CHO或CPAC的候选小区,并指示CHO或CPAC的执行条件与第一AI模型的第二输出相关联。

[0105] 在一些实施例中,第二输出指示以下至少一项:候选小区中的第一候选小区,或者终端设备对第一候选小区执行CHO或CPAC的概率。

[0106] 在一些实施例中,网络设备120可以向终端设备发送关于第一AI模型的有效性区域的第三信息。

[0107] 在一些实施例中,有效性区域包括以下至少一项:小区,无线电接入网通知区域或跟踪区域。

[0108] 在一些实施例中,网络设备120可以向终端设备发送对反馈信息的请求,并且网络设备120可以基于该请求来接收反馈信息。

[0109] 在一些实施例中,网络设备120可以基于反馈信息的预配置来接收关于第一AI模型的反馈信息。

[0110] 在一些实施例中,该反馈信息包括以下至少一项:在第一AI模型被使能时关于终端设备的移动性历史信息,关于用于空闲状态或非活动状态下的测量的功率的信息,在第一AI模型被使能时与向小区的无线电资源控制(RRC)建立失败或RRC恢复失败相关的信息,或者与在第一AI模型被使能时终端设备完成向小区的RRC建立过程或RRC恢复过程之后不久执行切换的情况有关的信息。

[0111] 在一些实施例中,网络设备120可以向终端设备发送关于第一AI模型的有效性定时器的第四信息。

[0112] 在一些实施例中,网络设备120可以向终端设备发送用于无线电承载或逻辑信道的服务质量(QoS)参数。

[0113] 在一些实施例中,该QoS参数包括用于无线电承载或逻辑信道的以下至少一项:数据包延迟预算,最大包错误率或丢包率,保证比特率,最大位速率,优先比特率,优先等级,

存活时间,或第五代QoS标识符值。

[0114] 图6是适合于实现本公开的一些实施例的设备600的简化框图。设备600可以被认为是如图1所示的终端设备110或网络设备120的另一示例实施例。因此,设备600可以在终端设备110或网络设备120处实现或实现为终端设备110或网络设备120的至少一部分。

[0115] 如图所示,设备600包括处理器610,耦合到处理器610的存储器620,耦合到处理器610的合适的发射机(TX)和接收机(RX)640,以及耦合到TX/RX640的通信接口。存储器620存储程序630的至少一部分。TX/RX640用于双向通信。TX/RX640具有至少一个天线以便于通信,尽管实际上本申请中提到的接入节点可以具有若干天线。通信接口可以表示与其他网络元件通信所需的任何接口,诸如用于gNB或eNB之间的双向通信的X2接口,用于移动性管理实体(MME)/服务网关(S-GW)与gNB或eNB之间的通信的S1接口,用于gNB或eNB与中继节点(RN)之间的通信的Un接口,或用于gNB或eNB与终端设备之间的通信的Uu接口。

[0116] 假定程序630包括程序指令,当由相关联的处理器610执行时,该程序指令使得设备600能够根据本公开的实施例进行操作,如这里参考图1至图5所讨论的。这里的实施例可以通过可由设备600的处理器610执行的计算机软件,或者通过硬件,或者通过软件和硬件的组合来实现。处理器610可经配置以实施本发明的各种实施例。此外,处理器610和存储器620的组合可以形成适于实现本公开的各种实施例的处理装置650。

[0117] 存储器620可以是适合于本地技术网络的任何类型,并且可以使用任何合适的数据存储技术来实现,作为非限制性示例,诸如非瞬态计算机可读存储介质,基于半导体的存储器设备,磁存储器设备和系统,光存储器设备和系统,固定存储器和可移动存储器。虽然在设备600中仅图示了一个存储器620,但是在设备600中可以有几个物理上不同的存储器模块。作为非限制性示例,处理器610可以是适合于本地技术网络的任何类型,并且可以包括通用计算机,专用计算机,微处理器,数字信号处理器(DSP)和基于多核处理器架构的处理器中的一个或多个。装置600可具有多个处理器,例如在时间上从属于使主处理器同步的时钟的专用集成电路芯片。

[0118] 包括在本公开的装置和/或设备中的组件可以以各种方式实现,包括软件,硬件,固件或其任何组合。在一个实施例中,可以使用例如存储在存储介质上的机器可执行指令的软件和/或固件来实现一个或多个单元。除了机器可执行指令之外或代替机器可执行指令,装置和/或设备中的部分或所有单元可以至少部分地由一个或多个硬件逻辑组件来实现。例如但不限于,可使用的硬件逻辑组件的说明性类型包括现场可编程门阵列(FPGA),专用集成电路(ASIC),专用标准产品(ASSP),片上系统系统(SOC),复杂可编程逻辑器件(CPLD)等。

[0119] 通常,本公开的各种实施例可以在硬件或专用电路,软件,逻辑或其任何组合中实现。一些方面可以用硬件来实现,而其他方面可以用固件或软件来实现,这些固件或软件可以由控制器,微处理器或其他计算设备来执行。虽然本公开的实施例的各方面被示出并描述为框图,流程图或使用一些其它图形表示,但将理解,本文描述的框,装置,系统,技术或方法可在作为非限制性示例的硬件,软件,固件,专用电路或逻辑,通用硬件或控制器或其它计算设备或其一些组合中实现。

[0120] 本公开还提供了有形地存储在非瞬态计算机可读存储介质上的至少一种计算机程序产品。计算机程序产品包括计算机可执行指令,例如包括在程序模块中的那些指令。

[0121] 用于执行本公开的方法的程序代码可以用一种或多种编程语言的任意组合来编写。这些程序代码可以被提供给通用计算机,专用计算机或其它可编程数据处理设备的处理器或控制器,使得程序代码在被处理器或控制器执行时使得流程图和/或框图中指定的功能/操作被实现。程序代码可以完全在机器上,部分在机器上,作为独立软件包,部分在机器上,部分在远程机器上或完全在远程机器或服务器上执行。

[0122] 上述程序代码可以包含在机器可读介质上,该机器可读介质可以是包含或存储由指令执行系统,装置或设备使用或结合指令执行系统,装置或设备使用的程序的任何有形介质。机器可读介质可以是机器可读信号介质或机器可读存储介质。机器可读介质可以包括但不限于电,磁,光,电磁,红外或半导体系统,装置或设备,或前述的任何合适的组合。机器可读存储介质的更具体的示例将包括具有一条或多条导线的电连接,便携式计算机磁盘,硬盘,随机存取存储器(RAM),只读存储器(ROM),可擦除可编程只读存储器(EPROM或闪存),光纤,便携式光盘只读存储器(CD-ROM),光存储装置,磁存储装置或前述的任何合适的组合。

[0123] 此外,虽然以特定顺序描述了操作,但是这不应被理解为要求以所示的特定顺序或按顺序执行这些操作,或者执行所有示出的操作,以实现期望的结果。在某些情况下,多任务和并行处理可能是有利的。同样地,虽然在上述讨论中包含了若干特定实施例细节,但是这些细节不应当被解释为对本公开的范围的限制,而应当被解释为对特定实施例特定的特征的描述。在单独实施例的上下文中描述的某些特征也可以在单个实施例中组合实现。相反,在单个实施例的上下文中描述的各种特征也可以在多个实施例中单独地或以任何合适的子组合来实现。

[0124] 尽管已经用结构特征和/或方法动作专用的语言描述了本公开,但是应当理解,所附权利要求中限定的本公开不必限于上述具体特征或动作。相反,上述具体特征和动作是作为实现权利要求的示例形式而公开的。

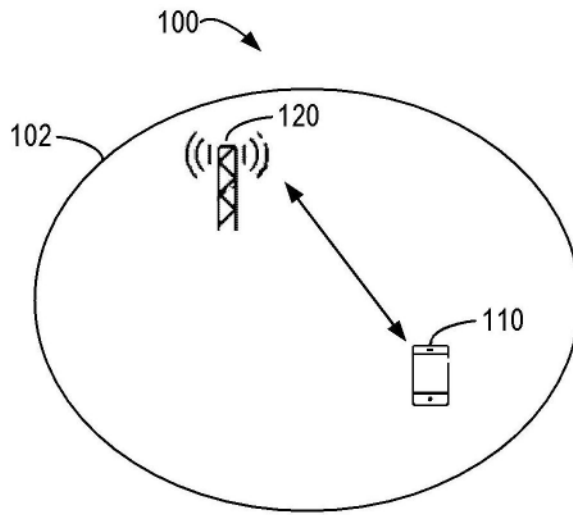


图1

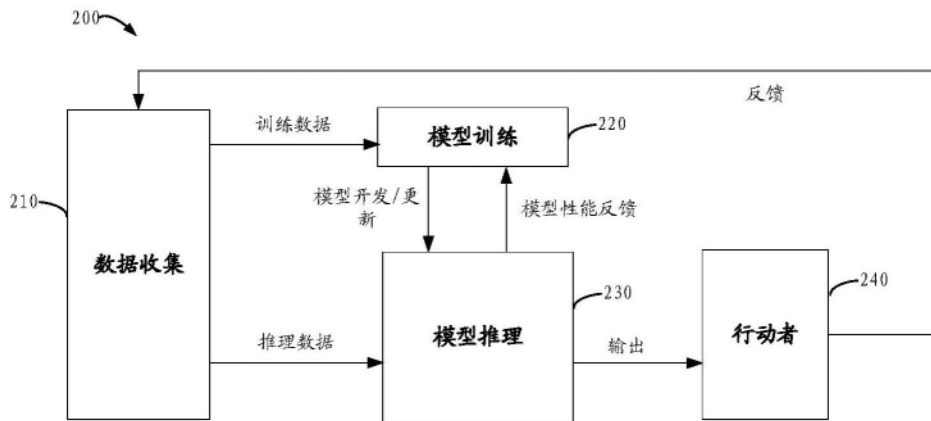


图2

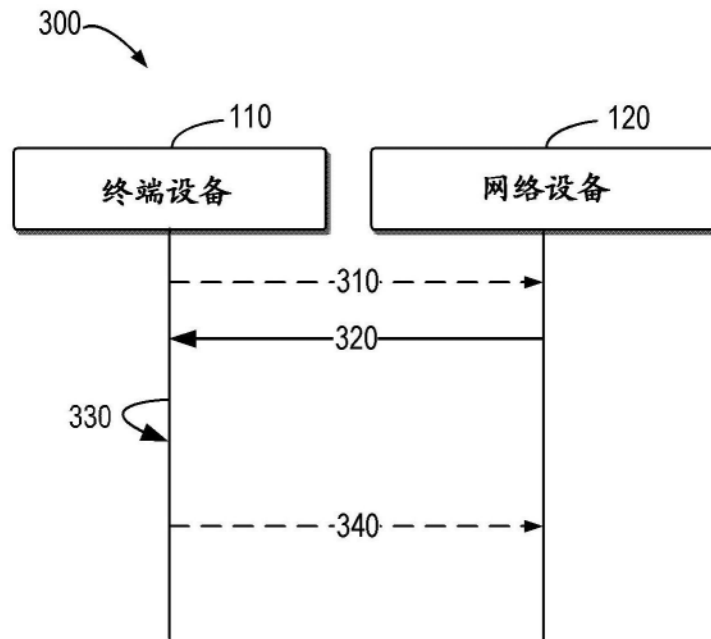


图3

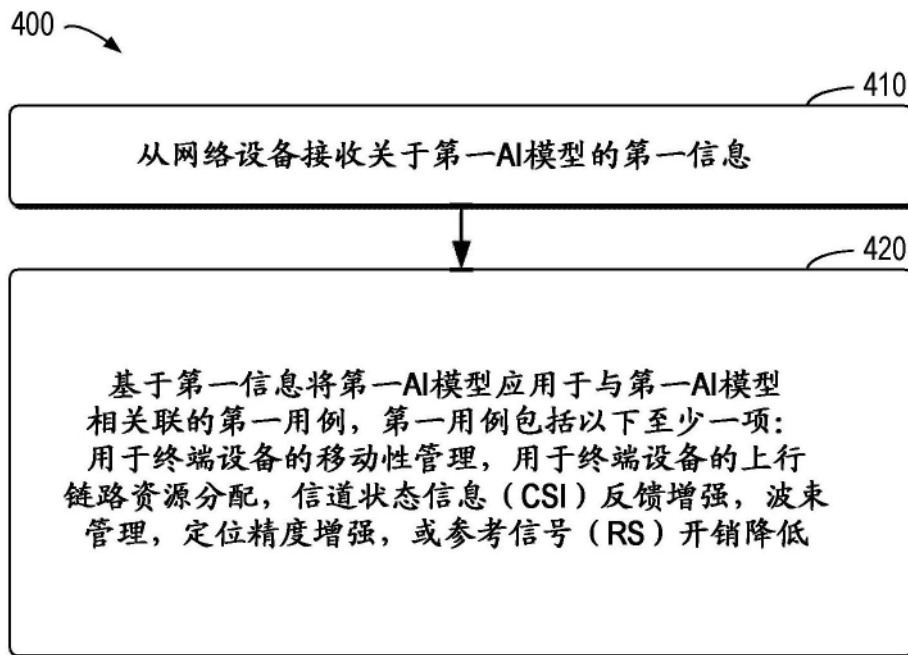


图4

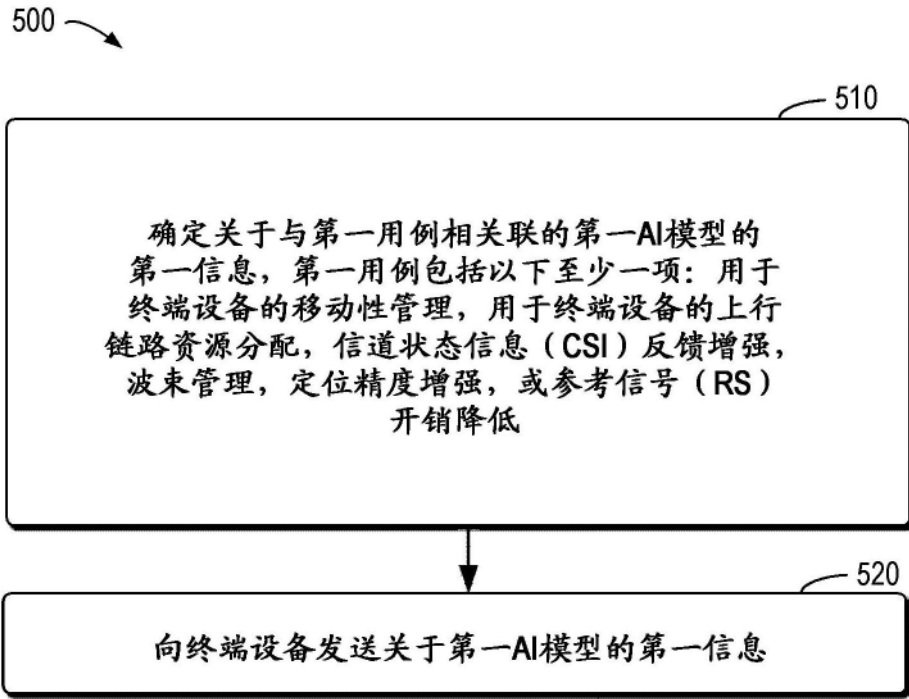


图5

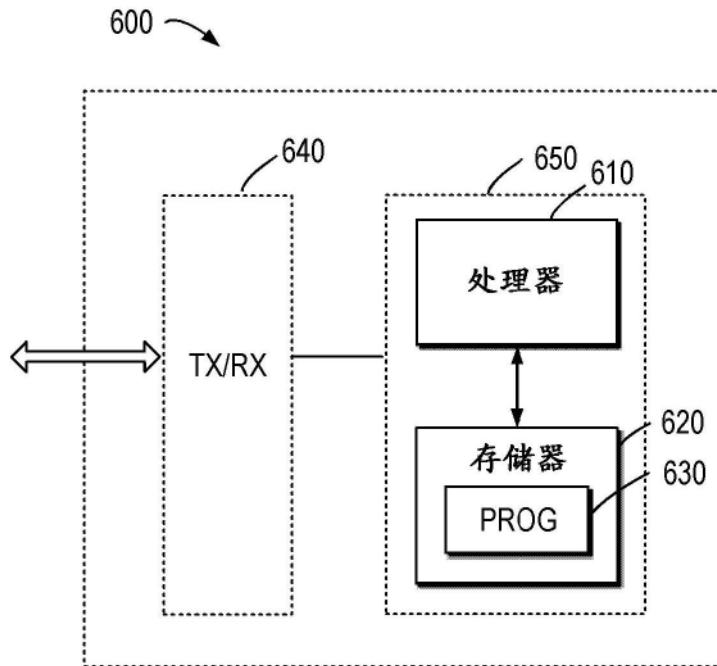


图6