

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2025 年 6 月 19 日 (19.06.2025)



(10) 国际公布号
WO 2025/124306 A1

- (51) 国际专利分类号:
H04W 16/18 (2009.01)
- (21) 国际申请号: PCT/CN2024/137406
- (22) 国际申请日: 2024 年 12 月 6 日 (06.12.2024)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202311733048.9 2023年12月15日 (15.12.2023) CN
- (71) 申请人: 中国移动通信有限公司研究院(CHINA MOBILE COMMUNICATION CO., LTD RESEARCH INSTITUTE) [CN/CN]; 中国北京市西城区宣武门西大街 32 号 100053 (CN)。中国移动通信集团有限公司(CHINA MOBILE COMMUNICATIONS GROUP CO., LTD.) [CN/CN]; 中国北京市西城区金融大街 29 号 100032 (CN)。
- (72) 发明人: 燕艺薇(YAN, Yiwei); 中国北京市西城区金融大街 29 号 100032 (CN)。孙奇(SUN, Qi); 中国

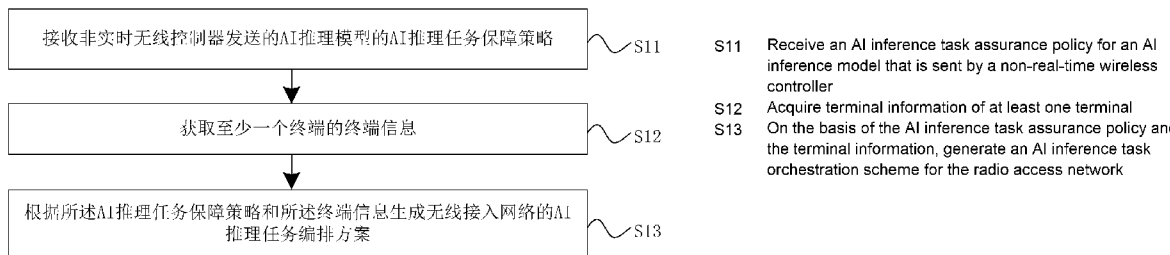
北京市西城区金融大街 29 号 100032 (CN)。李响(LI, Xiang); 中国北京市西城区金融大街 29 号 100032 (CN)。黄翊轩(HUANG, Yixuan); 中国北京市西城区金融大街 29 号 100032 (CN)。

(74) 代理人: 北京银龙知识产权代理有限公司(DRAGON INTELLECTUAL PROPERTY LAW FIRM); 中国北京市海淀区西直门北大街 32 号院枫蓝国际中心 2 号楼 10 层 100082 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(54) Title: AI INFERENCE TASK ORCHESTRATION METHOD FOR RADIO ACCESS NETWORK, APPARATUS, DEVICE AND STORAGE MEDIUM

(54) 发明名称: 无线接入网络的AI推理任务编排方法、装置、设备和存储介质



(57) Abstract: Provided in the present disclosure are an artificial intelligence (AI) inference task orchestration method for a radio access network, an apparatus, a device and a storage medium. AI technology is introduced to an RAN side. The method comprises: receiving an AI inference task assurance policy for an AI inference model that is sent by a non-real-time wireless controller, and acquiring terminal information of at least one terminal, the AI inference task assurance policy defining an optimization objective for the radio access network and feature information of the AI inference model, and the terminal information comprising computing resource information and communication resource information; and, on the basis of the AI inference task assurance policy and the terminal information, generating an AI inference task orchestration scheme for the radio access network.

(57) 摘要: 本公开提供了一种无线接入网络的AI推理任务编排方法、装置、设备和存储介质, 在RAN侧引入人工智能技术, 接收非实时无线控制器发送的AI推理模型的AI推理任务保障策略, 以及获取至少一个终端的终端信息, 其中, AI推理任务保障策略中定义了无线接入网络的优化目标和AI推理模型的特征信息, 终端信息包括计算资源信息和通信资源信息; 根据所述AI推理任务保障策略和所述终端信息生成无线接入网络的AI推理任务编排方案。

(84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚
(AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE,
BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR,
HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO,
PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN,
TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

无线接入网络的 AI 推理任务编排方法、装置、设备和存储介质

相关申请的交叉引用

本申请主张在 2023 年 12 月 15 日在中国提交的中国专利申请号 No. 202311733048.9 的优先权，其全部内容通过引用包含于此。

技术领域

本公开涉及无线通信技术领域，尤其涉及一种无线接入网络的 AI 推理任务编排方法、装置、设备和存储介质。

背景技术

随着无线局域网中计算能力开放研究的不断深入，基站作为一种边缘计算平台越来越受到业界的重视，无线局域网正朝着通信与计算一体化的方向发展。虽然现有研究提供了设备边缘计算协作中平衡精度和延迟的解决方案，但这些研究主要集中在边缘服务器上，并不能直接应用于基站与终端之间的计算协作，而基站中的计算资源是由通信处理和计算任务共享的，两者存在竞争关系。因此，在计算资源和通信的带宽资源受限的情况下，无线接入网络(Radio Access Network, RAN)侧需要根据动态变化的信道状态和不同任务的不同信息，对多基站多终端任务做近实时自适应的编排。

传统开放无线接入网络(Open-Radio Access Network, O-RAN)中的近实时控制器(Near-Real-Time RAN Intelligent Controller, Near-RT RIC)可以通过标准化 E2 接口采集近实时的通信相关的参数对系统性能进行智能化优化，然而这仅仅是对于 RAN 侧的通信相关参数的采集优化，缺乏对于计算资源和通信资源的统一感知调度能力，导致 RAN 侧的资源分配和协作计算效率低下，从而导致系统性能下降。

发明内容

本公开实施例的目的是提供一种无线接入网络的 AI 推理任务编排方法，在 RAN 侧引入人工智能(Artificial Intelligence, AI)技术，能根据实时的通信状态和计算能力，对 RAN 侧的计算节点做近实时的 AI 推理任务协作编排，提高了 RAN 侧资源分配和协作计算的效率，进而提升了系统性能。

为实现上述目的，本公开实施例提供了一种无线接入网络的 AI 推理任务编排方法，应用于近实时无线控制器，所述方法包括：

接收非实时无线控制器发送的 AI 推理模型的 AI 推理任务保障策略；其中，所述 AI 推理模型部署在无线接入网络中，所述 AI 推理任务保障策略中定义了无线接入网络的优化目标和

AI 推理模型的特征信息；

获取至少一个终端的终端信息；其中，所述终端信息包括计算资源信息和通信资源信息；根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案。

作为上述方案的改进，所述根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案，包括：

按照所述无线接入网络的优化目标，根据每一终端信息确定待分配节点和对应的计算节点，以及根据每一终端信息从所述 AI 推理模型的特征信息中选择对应的切分点和退出点；其中，所述切分点和所述退出点为所述 AI 推理模型中的其中一层；

以所述待分配节点及其对应的计算节点、所述切分点和所述退出点组成所述 AI 推理任务编排方案。

作为上述方案的改进，在根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案后，所述方法还包括：

将所述 AI 推理任务编排方案发送给待分配节点和对应的计算节点，以使所述待分配节点和对应的计算节点完成各自的推理计算任务。

作为上述方案的改进，当所述终端为用户设备时，所述获取至少一个终端的终端信息，包括：

向用户设备所驻留的基站发送请求消息，以使所述基站根据所述请求消息上报所述用户设备的终端信息；或，

向用户设备所驻留的基站订阅所述终端信息，以使所述基站上报所述用户设备的终端信息。

作为上述方案的改进，当所述终端为用户设备时，所述用户设备支持与近实时无线控制器的信息交互；则，所述获取至少一个终端的终端信息，包括：

向所述用户设备发送请求消息，以使所述用户设备根据所述请求消息上报终端信息；或，

向所述用户设备订阅所述终端信息，以使所述用户设备上报终端信息。

为实现上述目的，本公开实施例还提供了一种无线接入网络的 AI 推理任务编排方法，应用于非实时无线控制器，所述方法包括：

获取 AI 推理模型的模型信息；

根据所述模型信息生成 AI 推理任务保障策略；其中，所述 AI 推理模型部署在无线接入网络中，所述 AI 推理任务保障策略中定义了无线接入网络的优化目标和 AI 推理模型的特征信息；

将所述 AI 推理任务保障策略发送给近实时无线控制器，以使所述近实时无线控制器根据所述 AI 推理任务保障策略和终端信息生成无线接入网络的 AI 推理任务编排方案；其中，所述终端信息包括计算资源信息和通信资源信息。

作为上述方案的改进，所述模型信息包括所述 AI 推理模型的特征信息和 AI 推理任务的

性能保障参数；其中，所述性能保障参数包括但不限于模型推理准确度和单位时间的推理计算次数。

为实现上述目的，本公开实施例还提供了一种无线接入网络的 AI 推理任务编排装置，包括：

AI 推理任务保障策略接收模块，用于接收非实时无线控制器发送的 AI 推理模型的 AI 推理任务保障策略；其中，所述 AI 推理模型部署在无线接入网络中，所述 AI 推理任务保障策略中定义了无线接入网络的优化目标和 AI 推理模型的特征信息；

终端信息获取模块，用于获取至少一个终端的终端信息；其中，所述终端信息包括计算资源信息和通信资源信息；

AI 推理任务编排方案生成模块，用于根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案。

为实现上述目的，本公开实施例还提供了一种无线接入网络的 AI 推理任务编排装置，包括：

模型信息获取模块，用于获取 AI 推理模型的模型信息；

AI 推理任务保障策略生成模块，用于根据所述模型信息生成 AI 推理任务保障策略；其中，所述 AI 推理模型部署在无线接入网络中，所述 AI 推理任务保障策略中定义了无线接入网络的优化目标和 AI 推理模型的特征信息；

AI 推理任务保障策略发送模块，用于将所述 AI 推理任务保障策略发送给近实时无线控制器，以使所述近实时无线控制器根据所述 AI 推理任务保障策略和终端信息生成无线接入网络的 AI 推理任务编排方案；其中，所述终端信息包括计算资源信息和通信资源信息。

为实现上述目的，本公开实施例还提供了一种电子设备，包括处理器、存储器以及存储在所述存储器中且被配置为由所述处理器执行的计算机程序，所述处理器执行所述计算机程序时实现如上述任一实施例所述的无线接入网络的 AI 推理任务编排方法。

为实现上述目的，本公开实施例还提供了一种计算机可读存储介质，所述计算机可读存储介质包括存储的计算机程序，其中，在所述计算机程序运行时控制所述计算机可读存储介质所在设备执行如上述任一实施例所述的无线接入网络的 AI 推理任务编排方法。

相比于相关技术，本公开的无线接入网络的 AI 推理任务编排方法、装置、设备和存储介质，在 RAN 侧引入人工智能技术，首先接收非实时无线控制器发送的 AI 推理模型的 AI 推理任务保障策略，以及获取至少一个终端的终端信息，其中，AI 推理任务保障策略中定义了无线接入网络的优化目标和 AI 推理模型的特征信息，终端信息包括计算资源信息和通信资源信息；然后根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案，能根据实时的通信状态和计算能力，对 RAN 侧的计算节点做近实时的 AI 推理任务协作编排，提高了 RAN 侧资源分配和协作计算的效率，进而提升了系统性能。

附图说明

图 1 是本公开实施例提供的一种无线接入网络的 AI 推理任务编排方法的流程图；
图 2 是本公开实施例提供的近实时无线控制器和非实时无线控制器的信息交互示意图；
图 3 是本公开实施例提供的第一种用户设备的终端信息获取方式的流程图；
图 4 是本公开实施例提供的第二种用户设备的终端信息获取方式的流程图；
图 5 是本公开实施例提供的另一种无线接入网络的 AI 推理任务编排方法的流程图；
图 6 是本公开实施例提供的一种无线接入网络的 AI 推理任务编排装置的结构框图；
图 7 是本公开实施例提供的另一种无线接入网络的 AI 推理任务编排装置的结构框图；
图 8 是本公开实施例提供的一种电子设备的结构框图。

具体实施方式

下面将结合本公开实施例中的附图,对本公开实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本公开一部分实施例,而不是全部的实施例。基于本公开中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本公开保护的范围。

随着拓展现实(Extended Reality, XR)、自动驾驶、工业智能控制等新兴应用的兴起,深度学习相关算法例如图像识别等计算密集型任务受到更多关注。而移动设备的能力有限,无法满足其严格的计算要求和延迟要求,因此可借助 AI 边缘计算方式协作完成计算任务。AI 边缘计算方式可通过构建深度神经网络模型实现,由于深度神经网络模型结构中的卷积层是内核在输入张量的空间维度上进行点乘操作以生成输出张量的特征图,因此卷积层可以作为模型的切分点,由多个计算节点协同完成模型的推理。同时,还可以通过在模型结构中添加分支分类器,牺牲准确度的条件下提前退出推理以减少资源浪费。基于早期退出机制和模型分割技术,目前有许多针对边缘计算节点的协作推理的研究,通过为 AI 推理任务选择合适切分点和退出点,在计算资源和系统带宽的限制下满足时延和推理的精度要求。因此,本公开实施例中在 RAN 侧引入 AI 技术,协助终端完成计算任务。

参见图 1,图 1 是本公开实施例提供的一种无线接入网络的 AI 推理任务编排方法的流程图,所述无线接入网络的 AI 推理任务编排方法由近实时无线控制器执行实现,所述方法包括:

S11、接收非实时无线控制器发送的 AI 推理模型的 AI 推理任务保障策略;

S12、获取至少一个终端的终端信息;

S13、根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案。

值得说明的是,O-RAN 包括近实时控制器(Near-Real-Time RAN Intelligent Controller, Near-RT RIC)和非实时控制器(Non-Real-Time RAN Intelligent Controller, Non-RT RIC),所述终端包括但不限于基站、集中式单元、分布式单元(Distributed Unit, DU)和用户设备(User

Equipment, UE)，所述集中式单元包括（Centralized Unit, CU）、集中单元控制面（Centralized Unit-Control Plane, CU-CP）、分布单元用户面（Centralized Unit-User Plane, CU-UP）。参见图 2，图 2 是本公开实施例提供的近实时无线控制器和非实时无线控制器的信息交互示意图，所述非实时控制器通过开放和标准化的 A1 接口连接所述近实时控制器，所述非实时控制器的目的在于提供对应的机器学习模型来支持 RAN 智能化，以及由所述非实时控制器为所述近实时控制器提供 AI 推理模型和数据，在 O-RAN 架构中由于实时性的要求，所述近实时控制器在提供相应的功能时，通过利用已有的 AI 推理模型执行相关运算，所述 AI 推理模型能够进行故障告警分析、覆盖优化、参数优化、频谱分析、站间协同、移动性管理、切片管理、无线定位和环境感知识别中的一种或多种计算。所述近实时控制器和终端通过 E2 接口进行信息交互，本公开实施例中增强了 O-RAN 架构中的 E2 接口和 A1 接口功能，通过 A1 接口将非实时控制器根据 AI 推理模型的模型信息生成的 AI 推理任务保障策略下发给近实时控制器，再由 E2 接口采集实时的终端信息，根据终端信息和 AI 推理任务保障策略做近实时的 AI 推理任务协作编排，提高了 RAN 侧资源分配和协作计算效率，提升了系统性能。

具体地，在步骤 S11 中，所述非实时无线控制器根据 AI 推理模型的模型信息生成所述 AI 推理任务保障策略；其中，所述模型信息包括所述 AI 推理模型的特征信息和 AI 推理任务的性能保障参数。

示例性的，所述模型信息由所述非实时无线控制器向服务管理和流程编排框架（Service Management and Orchestration, SMO）获取得到，而对于 SMO 中所述模型信息的来源，可以由运营商直接配置或者与外部应用交互获取，当通过外部应用交互获取所述模型信息时，可预先设计相关接口或应用程序接口（Application Programming Interface, API），通过授权和认证机制与外部应用建立连接，以此获取所述模型信息。

示例性的，所述 AI 推理模型的特征信息包括 AI 推理模型标识（IDentifier, ID）、模型层数、各层类型、各层输出数据量、各层计算量、切分点位置、退出点位置、退出点准确度和退出点分类器计算量等。其中，每一所述特征信息对应的含义可参考表 1，具体示例可参考表 2。

表 1 特征参数及其对应含义

参数	含义
AI推理模型ID	指代 AI 推理模型的名称，例如 ResNet-18、LeNet 等。
模型层数	指该 AI 推理模型中包含的层的数量。
各层类型	数据类型为 array，指各层的类型，例如卷积层、池化层等。
各层输出数据量	数据类型为 array，指各层在数据流经过时产生的输出数据的数据量。
各层计算量	数据类型为 array，指各层在进行计算时所需的运算次数。
切分点位置	数据类型为 array，指将整个模型拆分成多个子模型时，可以划

	分的位置。
退出点位置	数据类型为 array，指在推理过程中，模型可以作为最终输出的位置。
退出点准确度	数据类型为 array，指在推理过程中，模型可以作为最终输出的位置后，所对应的准确度。
退出点分类器计算量	数据类型为 array，指在推理过程中，模型可以作为最终输出的位置后，所对应的分类器计算量。

表 2 ResNet-18 模型的特征信息示例

模型层数	各层类型	各层输出数据量	各层计算量	退出点准确度
1	conv1	25690112	118013952	
	Max pool	6422528	7225344	
2	conv2-1	6422528	115605504	
3	conv2-2	6422528	115605504	
4	conv2-3	6422528	115605504	
5 (Exit1)	conv2-4	6422528	115605504	63.91%
6	conv3-1	3211264	57802752	
7	conv3-2	3211264	115605504	
8	conv3-3	3211264	115605504	
9 (Exit2)	conv3-4	3211264	115605504	65.36%
10	conv4-1	1605632	57802752	
11	conv4-2	1605632	115605504	
12	conv4-3	1605632	115605504	
13 (Exit3)	conv4-4	1605632	115605504	67.64%
14	conv5-1	802816	57802752	
15	conv5-2	802816	115605504	
16	conv5-3	802816	115605504	
17	conv5-4	802816	115605504	
	avg pool	16384	1229312	
18 (Exit4)	fc	32000	1023000	68.96%

在表 2 中，conv 表示卷积层，max pool 表示最大池化层，avg pool 表示平均池化层，fc 表示全连接层，Exit 表示退出点位置，ResNet-18 模型一共有四个退出点，分别为位于第 5 层的 Exit1、位于第 9 层的 Exit2、位于第 13 层的 Exit3 和位于第 18 层的 Exit4，剩余的其余层可作为切分点。所述退出点准确度可以评估在选择这一层作为退出点时，其分类准确度，表格中的

准确度仅为示例。

示例性的，所述性能保障参数包括但不限于模型推理准确度、单位时间的推理计算次数、模型推理往返时延和模型频谱效率。其中，每一所述性能保障参数对应的含义可参考表 3。

表 3 性能保障参数及其对应含义

参数	含义
模型推理准确度	表示模型在进行推理时的预测准确程度。准确性是衡量模型性能的重要指标，它表示模型对输入数据进行预测的精度和正确性。
单位时间的推理计算次数	表示在单位时间内，模型能够进行的推理次数。用来评估和衡量模型的处理能力和效率。
模型推理往返时延 (Round Trip Time, RTT)	表示从输入数据传输到模型推理完成再返回结果所经历的总时间延迟。包含了数据传输、推理计算等各个环节的时间。
模型频谱效率	表示模型在推理任务的空口传输速率与占用带宽的比值。可用于评估对通信资源的利用率。

进一步地，所述非实时无线控制器在接收到所述模型信息后，根据所述模型信息生成 AI 推理任务保障策略，所述 AI 推理任务保障策略中定义了无线接入网络的优化目标、可优化参数和 AI 推理模型的特征信息。所述无线接入网络的优化目标指的是当前系统中保障策略的整体优化目标，例如最大化系统中模型推理准确度、最小化系统中模型推理往返时延等；所述可优化参数指的是所述近实时无线控制器可以为 AI 推理任务保障做调节的参数，例如 AI 推理模型的切分点选择、退出点选择等。假设所述无线接入网络的优化目标为：最大化系统中模型推理准确度、最小化系统中模型推理往返时延，即在选择完所述切分点和所述退出点后，需要满足模型推理准确度最大，以及满足模型推理往返时延最小，可通过对每一切分点和退出点进行轮询计算得到。

具体地，在步骤 S12 中，所述终端信息包括但不限于：计算资源信息、通信资源信息、当前小区服务用户信息、终端对 AI 推理任务的偏好信息。其中，所述计算资源信息包括但不限于：基站和终端的中央处理器（Central Processing Unit, CPU）利用率、CPU 频率、CPU 核绑定状态、剩余 CPU 核数量、CPU/图形处理器（Graphics Processing Unit, GPU）/嵌入式神经网络处理器（Neural-network Processing Unit, NPU）每秒浮点运算次数（Floating Point Operations Per Second, FLOPS）、GPU 显存容量/剩余显存、GPU cuda 核心数/剩余核心数等；所述通信资源信息包括但不限于：基站的上/下行总带宽、剩余带宽、物理资源块（Physical Resource Block, PRB）数、PRB 利用率等，以及还可以包括终端的信道条件、本基站到邻站的通信链路时延等，例如终端的信道条件包括信噪比（Signal-to-Noise Ratio, SNR）、接收的信号强度指示（Received Signal Strength Indication, RSSI）等；所述当前小区服务用户信息包括但不限于：用户类型标识（非 AI 用户或 AI 用户）、AI 推理任务模型标识（例如：ResNet-18、视觉几何组（Visual

Geometry Group, VGG) 等) 等; 所述终端对 AI 推理任务的偏好信息包括但不限于: 对每种 AI 推理任务的切分点选择偏好等。

值得说明的是, 当所述终端为基站或所述集中式单元和所述分布式单元所在基站时, 所述终端信息可由基站直接发送给所述近实时无线控制器。当所述终端为用户设备时, 本公开实施例中提供两种获取用户设备的终端信息的方式, 第一种是通过基站获取, 第二种是直接通过用户设备获取。本公开实施例所述的基站可能有多种形式, 比如宏基站、微基站、中继站或接入点等。该基站可以是一体化基站, 或者可以是包括集中式单元 CU 和分布式单元 DU 的基站。

在第一种实施方式中, 当所述终端为用户设备时, 所述获取至少一个终端的终端信息, 包括: 向用户设备所驻留的基站发送请求消息, 以使所述基站根据所述请求消息上报所述用户设备的终端信息; 或, 向用户设备所驻留的基站订阅所述终端信息, 以使所述基站上报所述用户设备的终端信息。

示例性的, 参见图 3, 近实时无线控制器通过 E2 接口向用户设备所驻留的基站订阅或请求所述终端信息, 可以是周期性订阅或事件性触发上报, 事件性触发包括但不限于: 基站算力波动、终端算力波动、基站带宽波动、终端电量波动、终端切分模式偏好变化等。基站根据需求采集用户设备的空口及终端数据, 汇总成终端信息, 并向近实时无线控制器上报所述用户设备的终端信息。另外, 所述基站还将其自身的终端信息同步上报给所述近实时无线控制器。

在第二种实施方式中, 当所述终端为用户设备时, 所述用户设备支持与近实时无线控制器的信息交互; 则, 所述获取至少一个终端的终端信息, 包括: 向所述用户设备发送请求消息, 以使所述用户设备根据所述请求消息上报终端信息; 或, 向所述用户设备订阅所述终端信息, 以使所述用户设备上报终端信息。

示例性的, 参见图 4, 所述用户设备与所述近实时无线控制器在现有协议栈基础上使得双方支持对应的接口应用层协议, 从而实现双方的信息交互。此时近实时无线控制器通过 E2 接口向用户设备订阅或请求所述终端信息, 可以是周期性订阅或事件性触发上报, 用户设备根据需求采集空口及终端数据, 汇总成终端信息, 并向近实时无线控制器上报所述终端信息。所述近实时无线控制器同步发送请求消息给所述基站或订阅基站的终端信息, 以使所述基站上报其自身的终端信息。

具体地, 在步骤 S13 中, 所述根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案, 包括:

S131、按照所述无线接入网络的优化目标, 根据每一终端信息确定待分配节点和对应的计算节点, 以及根据每一终端信息从所述 AI 推理模型的特征信息中选择对应的切分点和退出点; 其中, 所述计算节点为基站, 所述切分点和所述退出点为所述 AI 推理模型中的其中一层;

S132、以所述待分配节点及其对应的计算节点、所述切分点和所述退出点组成所述 AI 推理任务编排方案。

示例性的, 所述待分配节点为自身无法满足计算需求, 且需要借助计算节点分担推理计算

任务的终端；所述计算节点为可接收所述待分配节点的推理计算任务的终端。所述近实时无线控制器在接收到至少一个终端发送的终端信息后，确定待分配节点和计算节点，会应用优化算法（如差分算法），评估这些待分配节点所需的计算节点，以及进行切分点和退出点的选择。如根据所述计算资源信息确定其中一个待分配节点需要较大计算量时，可以分配较多的层数（即切分点和退出点之间的层数相隔较多）给这一待分配节点，反之则可以分配较少的层数；如根据所述通信资源信息可以确定这一待分配节点需要剩余带宽较大的计算节点时，优先分配满足条件的计算节点作为这一待分配节点的计算节点；如此时其中一个终端信息中给出了 AI 推理任务模型标识，则在生成 AI 推理任务编排方案时，会根据 AI 推理任务模型标识选择对应的 AI 推理模型进行推理任务；如此时其中一个终端信息中给出了对每种 AI 推理任务的切分点选择偏好，则在生成 AI 推理任务编排方案时，会优先根据这一待分配节点的切分点选择偏好去分配切分点。

具体地，在根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案后，所述方法还包括：

S14、将所述 AI 推理任务编排方案发送给待分配节点和对应的计算节点，以使所述待分配节点和对应的计算节点完成各自的推理计算任务。

示例性的，假设所述待分配节点为用户设备，所述计算节点为基站，此时收集到三个用户设备和两个基站的终端信息，两个基站和三个用户设备满足以下情况：基站 A 为用户设备 1 提供通信连接服务，基站 B 为用户设备 2、3 提供通信连接服务，基站 A 和 B 均提供计算服务，AI 推理模型皆为表 2 中的 ResNet-18。假设此时编排方案是为用户设备 1~3 选择合适的计算节点，并为 AI 推理任务选择其处理的切分点和退出点，AI 推理任务编排方案如下表 4 所示。

表 4 AI 推理任务编排方案示例

用户设备	计算节点	切分点	退出点
用户设备 1	基站 A	Layer ID = 3	Exit2
用户设备 2	基站 A	Layer ID = 8	Exit3
用户设备 3	基站 B	Layer ID = 11	Exit3

示例性的，按照表 4 以用户设备 1 为例进行说明，此时为用户设备 1 选择的计算节点为基站 A，即由基站 A 与用户设备 1 共同完成运算任务，切分点为第 3 层，退出点为 Exit2，则基站 A 完成的计算量为第 3 层到第 9 层的计算量，根据选择的切分点，由用户设备 1 完成前面几层的推理计算，切分点这一层生成的特征图传输给基站 A 后，由基站 A 完成从第 3 层到第 9 层的推理计算，最终将生成的结果传输回用户设备 1。因此，切分点和退出点的选择决定了待分配节点和计算节点各自的计算量，退出点的选择也决定了该模型的准确度。

相比于相关技术，本公开的无线接入网络的 AI 推理任务编排方法，在 RAN 侧引入人工智能技术，首先接收非实时无线控制器发送的 AI 推理模型的 AI 推理任务保障策略，以及获取至少一个终端的终端信息，其中，AI 推理任务保障策略中定义了无线接入网络的优化目标

和 AI 推理模型的特征信息,终端信息包括计算资源信息和通信资源信息;然后根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案,通过对计算节点、任务切分点和退出点的选择,能根据实时的通信状态和计算能力,对 RAN 侧的计算节点做近实时的 AI 推理任务协作编排,提高了 RAN 侧资源分配和协作计算的效率,进而提升了系统性能。

参见图 5,图 5 是本公开实施例提供的另一种无线接入网络的 AI 推理任务编排方法的流程图,所述无线接入网络的 AI 推理任务编排方法由非实时无线控制器执行实现,所述方法包括:

S21、获取 AI 推理模型的模型信息;

S22、根据所述模型信息生成 AI 推理任务保障策略;

S23、将所述 AI 推理任务保障策略发送给近实时无线控制器,以使所述近实时无线控制器根据所述 AI 推理任务保障策略和终端信息生成无线接入网络的 AI 推理任务编排方案。

具体地,所述模型信息包括所述 AI 推理模型的特征信息和 AI 推理任务的性能保障参数;其中,所述性能保障参数包括但不限于模型推理准确度和单位时间的推理计算次数。

值得说明的是,本公开实施例所述的无线接入网络的 AI 推理任务编排方法的具体工作过程可参考上述实施例,在此不再赘述。

参见图 6,图 6 是本公开实施例提供的一种无线接入网络的 AI 推理任务编排装置 100 的结构框图,所述无线接入网络的 AI 推理任务编排装置 100 包括:

AI 推理任务保障策略接收模块 11,用于接收非实时无线控制器发送的 AI 推理模型的 AI 推理任务保障策略;其中,所述 AI 推理模型部署在无线接入网络中,所述 AI 推理任务保障策略中定义了无线接入网络的优化目标和 AI 推理模型的特征信息;

终端信息获取模块 12,用于获取至少一个终端的终端信息;其中,所述终端信息包括计算资源信息和通信资源信息;

AI 推理任务编排方案生成模块 13,用于根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案。

具体地,所述 AI 推理任务编排方案生成模块 13 具体用于:

按照所述无线接入网络的优化目标,根据每一终端信息确定待分配节点和对应的计算节点,以及根据每一终端信息从所述 AI 推理模型的特征信息中选择对应的切分点和退出点;其中,所述计算节点为基站,所述切分点和所述退出点为所述 AI 推理模型中的其中一层;

以所述待分配节点及其对应的计算节点、所述切分点和所述退出点组成所述 AI 推理任务编排方案。

具体地,所述无线接入网络的 AI 推理任务编排装置 100 还包括:

AI 推理任务编排方案发送模块,用于将所述 AI 推理任务编排方案发送给待分配节点和对应的计算节点,以使所述待分配节点和对应的计算节点完成各自的推理计算任务。

具体地，当所述终端为用户设备时，所述终端信息获取模块 12 具体用于：向用户设备所驻留的基站发送请求消息，以使所述基站根据所述请求消息上报所述用户设备的终端信息；或，向用户设备所驻留的基站订阅所述终端信息，以使所述基站上报所述用户设备的终端信息。

具体地，当所述终端为用户设备时，所述用户设备支持与近实时无线控制器的信息交互；则，所述终端信息获取模块 12 具体用于：向所述用户设备发送请求消息，以使所述用户设备根据所述请求消息上报终端信息；或，向所述用户设备订阅所述终端信息，以使所述用户设备上报终端信息。

参见图 7，图 7 是本公开实施例提供的另一种无线接入网络的 AI 推理任务编排装置 200 的结构框图，所述无线接入网络的 AI 推理任务编排装置 200 包括：

模型信息获取模块 21，用于获取 AI 推理模型的模型信息；

AI 推理任务保障策略生成模块 22，用于根据所述模型信息生成 AI 推理任务保障策略；其中，所述 AI 推理模型部署在无线接入网络中，所述 AI 推理任务保障策略中定义了无线接入网络的优化目标和 AI 推理模型的特征信息；

AI 推理任务保障策略发送模块 23，用于将所述 AI 推理任务保障策略发送给近实时无线控制器，以使所述近实时无线控制器根据所述 AI 推理任务保障策略和终端信息生成无线接入网络的 AI 推理任务编排方案；其中，所述终端信息包括计算资源信息和通信资源信息。

具体地，所述模型信息包括所述 AI 推理模型的特征信息和 AI 推理任务的性能保障参数；其中，所述性能保障参数包括但不限于模型推理准确度和单位时间的推理计算次数。

相比于相关技术，本公开的无线接入网络的 AI 推理任务编排装置，在 RAN 侧引入人工智能技术，首先接收非实时无线控制器发送的 AI 推理模型的 AI 推理任务保障策略，以及获取至少一个终端的终端信息，其中，AI 推理任务保障策略中定义了无线接入网络的优化目标和 AI 推理模型的特征信息，终端信息包括计算资源信息和通信资源信息；然后根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案，通过对计算节点、任务切分点和退出点的选择，能根据实时的通信状态和计算能力，对 RAN 侧的计算节点做近实时的 AI 推理任务协作编排，提高了 RAN 侧资源分配和协作计算的效率，进而提升了系统性能。

参见图 8，图 8 是本公开实施例提供的一种电子设备 300 的结构框图，所述电子设备 300 包括处理器 31、存储器 32 以及存储在所述存储器 32 中并可在所述处理器 31 上运行的计算机程序。所述处理器 31 执行所述计算机程序时实现上述各个无线接入网络的 AI 推理任务编排方法实施例中的步骤，比如步骤 S11~S13、S21~S23。

示例性的，所述计算机程序可以被分割成一个或多个模块/单元，所述一个或者多个模块/单元被存储在所述存储器 32 中，并由所述处理器 31 执行，以完成本公开。所述一个或多个模块/单元可以是能够完成特定功能的一系列计算机程序指令段，该指令段用于描述所述计算机程序在所述电子设备 300 中的执行过程。

所述电子设备 300 可包括，但不限于，处理器 31、存储器 32。本领域技术人员可以理解，所述示意图仅仅是电子设备 300 的示例，并不构成对电子设备 300 的限定，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，例如所述电子设备 300 还可以包括输入输出设备、网络接入设备、总线等。

所述处理器 31 可以是中央处理单元（Central Processing Unit, CPU），还可以是其他通用处理器、数字信号处理器（Digital Signal Processor, DSP）、专用集成电路（Application Specific Integrated Circuit, ASIC）、现成可编程门阵列（Field-Programmable Gate Array, FPGA）或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等，所述处理器 31 是所述电子设备 300 的控制中心，利用各种接口和线路连接整个电子设备 300 的各个部分。

所述存储器 32 可用于存储所述计算机程序和/或模块，所述处理器 31 通过运行或执行存储在所述存储器 32 内的计算机程序和/或模块，以及调用存储在存储器 32 内的数据，实现所述电子设备 300 的各种功能。所述存储器 32 可主要包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序（比如声音播放功能、图像播放功能等）等；存储数据区可存储根据手机的使用所创建的数据（比如音频数据、电话本等）等。此外，存储器 32 可以包括高速随机存取存储器，还可以包括非易失性存储器，例如硬盘、内存、插接式硬盘，智能存储卡（Smart Media Card, SMC），安全数字（Secure Digital, SD）卡，闪存卡（Flash Card）、至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。

其中，所述电子设备 300 集成的模块/单元如果以软件功能单元的形式实现并作为独立的产品销售或使用，可以存储在一个计算机可读取存储介质中。基于这样的理解，本公开实现上述实施例方法中的全部或部分流程，也可以通过计算机程序来指令相关的硬件来完成，所述的计算机程序可存储于一计算机可读存储介质中，该计算机程序在被处理器 31 执行时，可实现上述各个方法实施例的步骤。其中，所述计算机程序包括计算机程序代码，所述计算机程序代码可以为源代码形式、对象代码形式、可执行文件或某些中间形式等。所述计算机可读介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U 盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器（Read-Only Memory, ROM）、随机存取存储器（Random Access Memory, RAM）、电载波信号、电信号以及软件分发介质等。

以上所述是本公开的优选实施方式，应当指出，对于本技术领域的普通技术人员来说，在不脱离本公开原理的前提下，还可以做出若干改进和润饰，这些改进和润饰也视为本公开的保护范围。

权利要求书

1、一种无线接入网络的 AI 推理任务编排方法,应用于近实时无线控制器,所述方法包括:

接收非实时无线控制器发送的 AI 推理模型的 AI 推理任务保障策略;其中,所述 AI 推理模型部署在无线接入网络中,所述 AI 推理任务保障策略中定义了无线接入网络的优化目标和 AI 推理模型的特征信息;

获取至少一个终端的终端信息;其中,所述终端信息包括计算资源信息和通信资源信息;

根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案。

2、如权利要求 1 所述的无线接入网络的 AI 推理任务编排方法,其中,所述根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案,包括:

按照所述无线接入网络的优化目标,根据每一终端信息确定待分配节点和对应的计算节点,以及根据每一终端信息从所述 AI 推理模型的特征信息中选择对应的切分点和退出点;其中,所述切分点和所述退出点为所述 AI 推理模型中的其中一层;

以所述待分配节点及其对应的计算节点、所述切分点和所述退出点组成所述 AI 推理任务编排方案。

3、如权利要求 2 所述的无线接入网络的 AI 推理任务编排方法,其中,在根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案后,所述方法还包括:

将所述 AI 推理任务编排方案发送给待分配节点和对应的计算节点,以使所述待分配节点和对应的计算节点完成各自的推理计算任务。

4、如权利要求 1 所述的无线接入网络的 AI 推理任务编排方法,其中,当所述终端为用户设备时,所述获取至少一个终端的终端信息,包括:

向用户设备所驻留的基站发送请求消息,以使所述基站根据所述请求消息上报所述用户设备的终端信息;或,

向用户设备所驻留的基站订阅所述终端信息,以使所述基站上报所述用户设备的终端信息。

5、如权利要求 1 所述的无线接入网络的 AI 推理任务编排方法,其中,当所述终端为用户设备时,所述用户设备支持与近实时无线控制器的信息交互;则,所述获取至少一个终端的终端信息,包括:

向所述用户设备发送请求消息,以使所述用户设备根据所述请求消息上报终端信息;或,向所述用户设备订阅所述终端信息,以使所述用户设备上报终端信息。

6、一种无线接入网络的 AI 推理任务编排方法,应用于非实时无线控制器,所述方法包括:

获取 AI 推理模型的模型信息;

根据所述模型信息生成 AI 推理任务保障策略;其中,所述 AI 推理模型部署在无线接入

网络中，所述 AI 推理任务保障策略中定义了无线接入网络的优化目标和 AI 推理模型的特征信息；

将所述 AI 推理任务保障策略发送给近实时无线控制器，以使所述近实时无线控制器根据所述 AI 推理任务保障策略和终端信息生成无线接入网络的 AI 推理任务编排方案；其中，所述终端信息包括计算资源信息和通信资源信息。

7、如权利要求 6 所述的无线接入网络的 AI 推理任务编排方法，其中，所述模型信息包括所述 AI 推理模型的特征信息和 AI 推理任务的性能保障参数；其中，所述性能保障参数包括但不限于模型推理准确度和单位时间的推理计算次数。

8、一种无线接入网络的 AI 推理任务编排装置，包括：

AI 推理任务保障策略接收模块，用于接收非实时无线控制器发送的 AI 推理模型的 AI 推理任务保障策略；其中，所述 AI 推理模型部署在无线接入网络中，所述 AI 推理任务保障策略中定义了无线接入网络的优化目标和 AI 推理模型的特征信息；

终端信息获取模块，用于获取至少一个终端的终端信息；其中，所述终端信息包括计算资源信息和通信资源信息；

AI 推理任务编排方案生成模块，用于根据所述 AI 推理任务保障策略和所述终端信息生成无线接入网络的 AI 推理任务编排方案。

9、一种无线接入网络的 AI 推理任务编排装置，包括：

模型信息获取模块，用于获取 AI 推理模型的模型信息；

AI 推理任务保障策略生成模块，用于根据所述模型信息生成 AI 推理任务保障策略；其中，所述 AI 推理模型部署在无线接入网络中，所述 AI 推理任务保障策略中定义了无线接入网络的优化目标和 AI 推理模型的特征信息；

AI 推理任务保障策略发送模块，用于将所述 AI 推理任务保障策略发送给近实时无线控制器，以使所述近实时无线控制器根据所述 AI 推理任务保障策略和终端信息生成无线接入网络的 AI 推理任务编排方案；其中，所述终端信息包括计算资源信息和通信资源信息。

10、一种电子设备，包括处理器、存储器以及存储在所述存储器中且被配置为由所述处理器执行的计算机程序，所述处理器执行所述计算机程序时实现如权利要求 1 至 7 中任意一项所述的无线接入网络的 AI 推理任务编排方法。

11、一种计算机可读存储介质，所述计算机可读存储介质包括存储的计算机程序，其中，在所述计算机程序运行时控制所述计算机可读存储介质所在设备执行如权利要求 1 至 7 中任意一项所述的无线接入网络的 AI 推理任务编排方法。

12、一种计算机程序产品，包括计算机指令，所述计算机指令被处理器执行时实现如权利要求 1 至 7 中任一项所述的方法中的步骤。

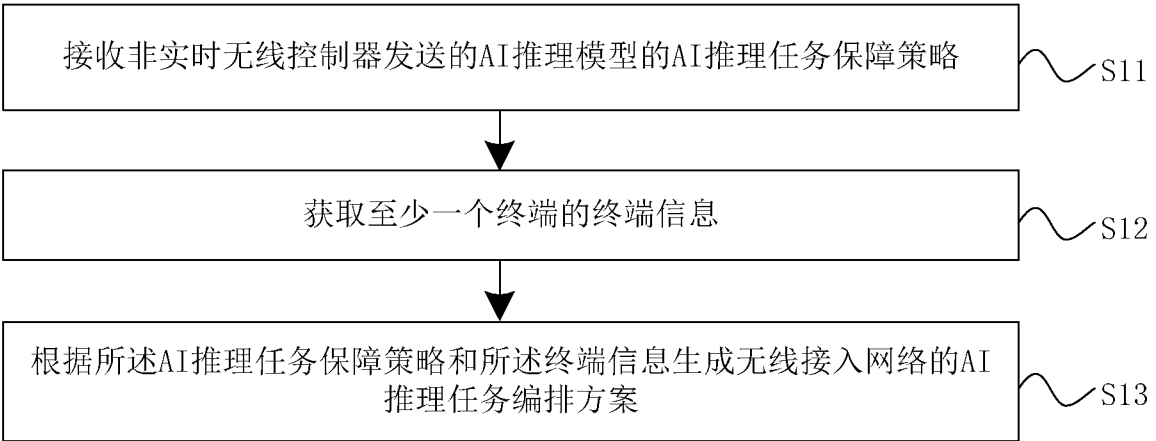


图 1

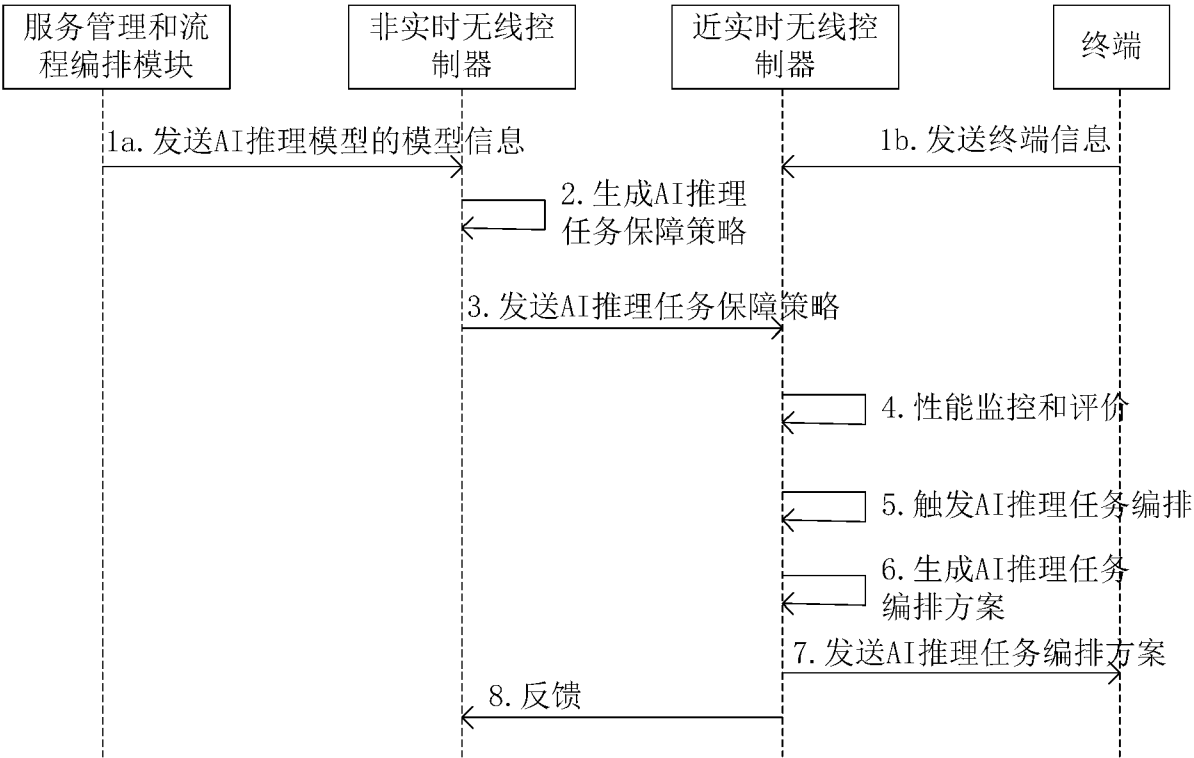


图 2

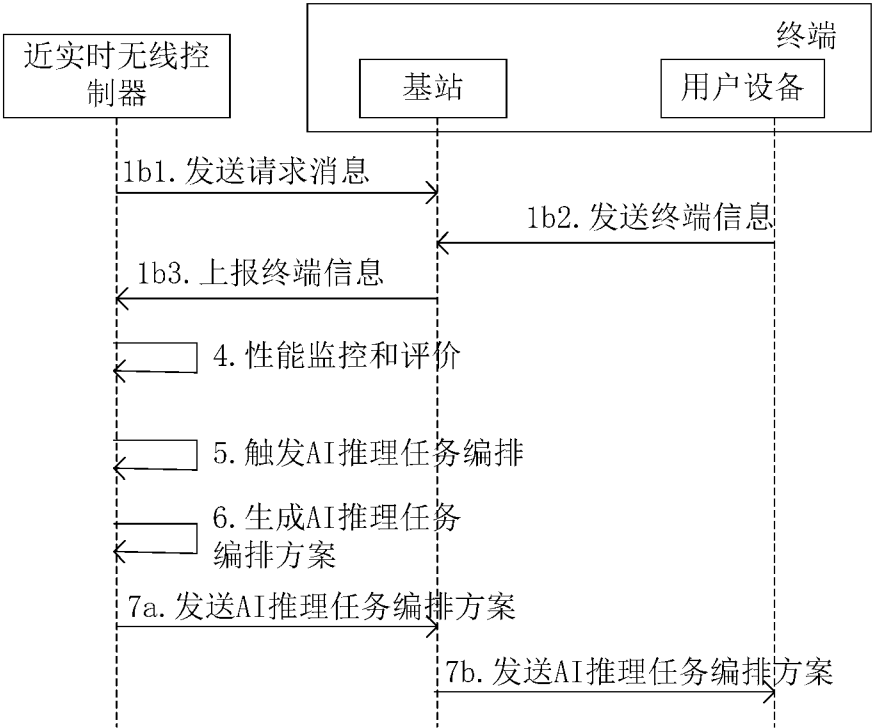


图 3

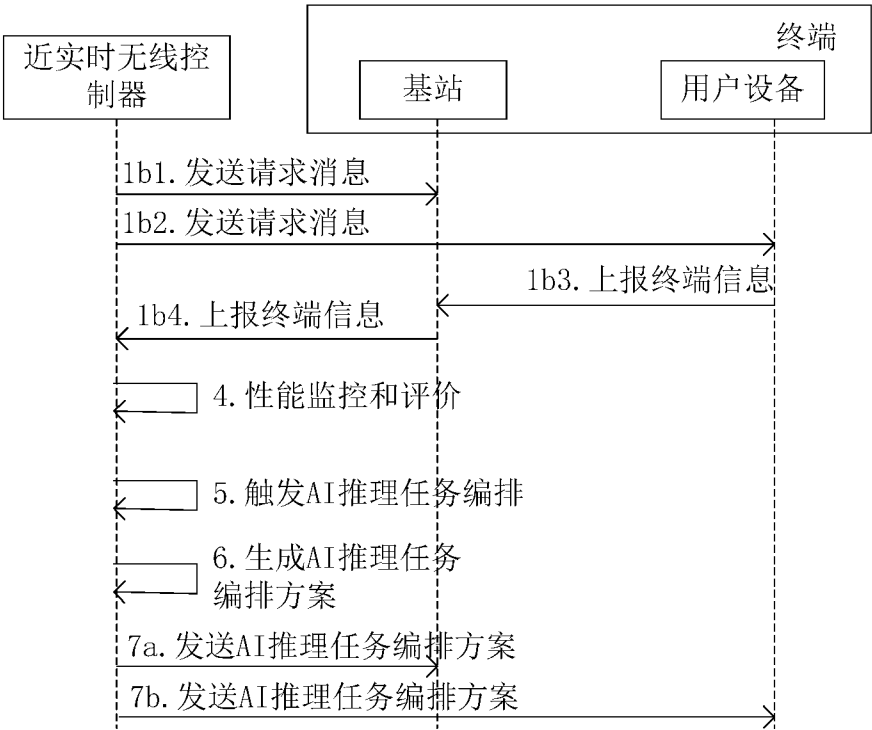


图 4

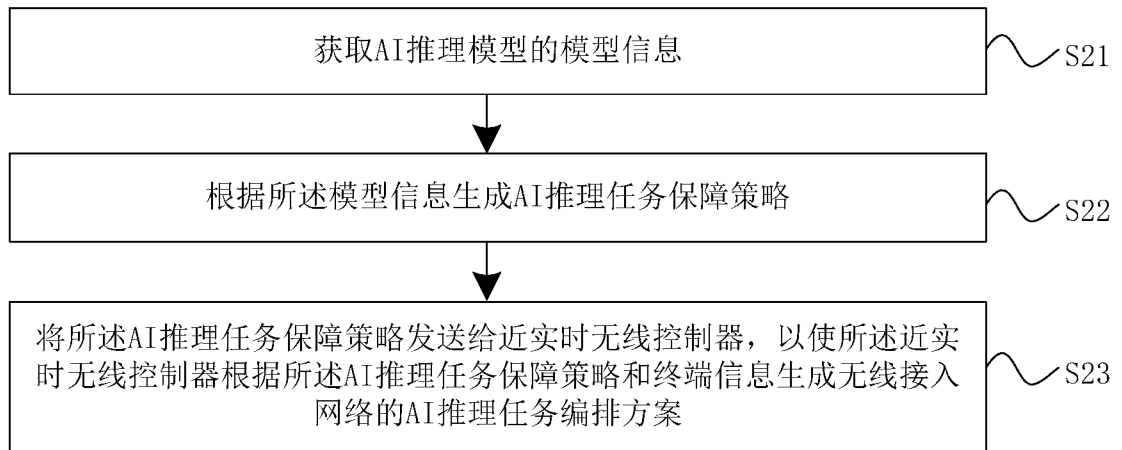


图 5

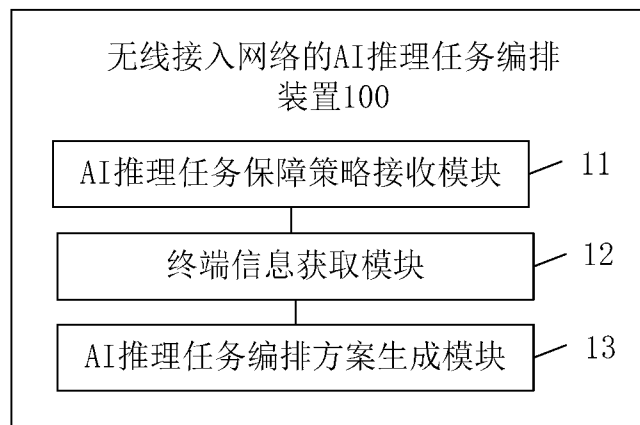


图 6

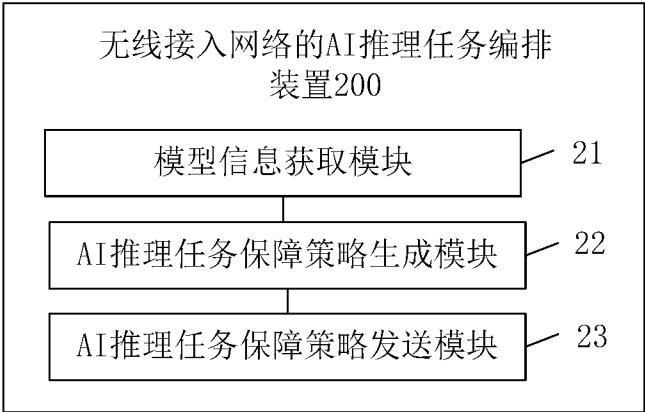


图 7

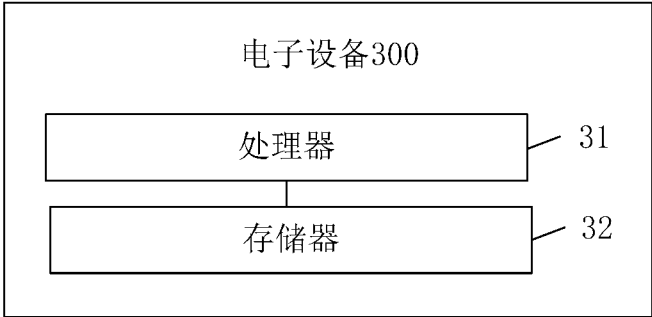


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/137406

A. CLASSIFICATION OF SUBJECT MATTER

H04W 16/18(2009.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04W, H04L, H04B, H04Q 7/-

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, CNKI, VEN, WOTXT, USTXT, EPTXT, 3GPP: 开放无线接入网络, 非实时, 准实时, 近实时, 无线网络控制器, 人工智能, 编排, 编排器, Open-Radio Access Network, O-RAN, non-real-time, near real-time, Non-RT, Near-RT, RAN Intelligent Controller, RIC, Artificial Intelligence, AI, Orchestration, Orchestrator

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	薛旭等 (XUE, Xu et al.). "开放智能无线网络架构和平台设计研究 (Exploring the Design of an Open Intelligent Wireless Network Architecture and Platform)" 移动通信 (Mobile Communications), No. 523, 11 April 2023 (2023-04-11), section 1	1, 4-12
A	CN 113875278 A (APPLE INC.) 31 December 2021 (2021-12-31) entire document	1-12
A	CN 113923694 A (PURPLE MOUNTAIN LABORATORIES FOR NETWORK AND COMMUNICATION SECURITY) 11 January 2022 (2022-01-11) entire document	1-12
A	CN 115988580 A (CHINA UNITED NETWORK COMMUNICATIONS GROUP CO., LTD.) 18 April 2023 (2023-04-18) entire document	1-12



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

14 March 2025

Date of mailing of the international search report

19 March 2025

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
China No. 6, Xitucheng Road, Jimenqiao, Haidian District,
Beijing 100088

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2024/137406

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	113875278	A	31 December 2021	US	2022159525	A1	19 May 2022
				WO	2020242987	A1	03 December 2020
CN	113923694	A	11 January 2022	None			
CN	115988580	A	18 April 2023	None			

A. 主题的分类

H04W 16/18(2009.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC: H04W, H04L, H04B, H04Q 7/-

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS, CNTXT, CNKI, VEN, WOTXT, USTXT, EPTXT, 3GPP: 开放无线接入网络, 非实时, 准实时, 近实时, 无线网络控制器, 人工智能, 编排, 编排器, Open-Radio Access Network, O-RAN, non-real-time, near real-time, Non-RT, Near-RT, RAN Intelligent Controller, RIC, Artificial Intelligence, AI, Orchestration, Orchestrator

C. 相关文件

类型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	薛旭等. "开放智能无线网络架构和平台设计研究" 移动通信, 第523期, 2023年4月11日 (2023 - 04 - 11), 第1节	1, 4-12
A	CN 113875278 A (苹果公司) 2021年12月31日 (2021 - 12 - 31) 全文	1-12
A	CN 113923694 A (网络通信与安全紫金山实验室) 2022年1月11日 (2022 - 01 - 11) 全文	1-12
A	CN 115988580 A (中国联合网络通信集团有限公司) 2023年4月18日 (2023 - 04 - 18) 全文	1-12

☐ 其余文件在C栏的续页中列出。

☒ 见同族专利附件。

* 引用文件的具体类型:

"A" 认为不特别相关的表示了现有技术一般状态的文件

"D" 申请人在国际申请中引证的文件

"E" 在国际申请日的当天或之后公布的在先申请或专利

"L" 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

"O" 涉及口头公开、使用、展览或其他方式公开的文件

"P" 公布日先于国际申请日但迟于所要求的优先权日的文件

"T" 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

"X" 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

"Y" 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

"&" 同族专利的文件

国际检索实际完成的日期

2025年3月14日

国际检索报告邮寄日期

2025年3月19日

ISA/CN的名称和邮寄地址

中国国家知识产权局
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

芦婧

电话号码 (+86) 010-62412158

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/137406

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	113875278	A	2021年12月31日	US	2022159525	A1	2022年5月19日
				WO	2020242987	A1	2020年12月3日
CN	113923694	A	2022年1月11日	无			
CN	115988580	A	2023年4月18日	无			