

(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 4 月 3 日 (03.04.2025)



(10) 国际公布号
WO 2025/065720 A1

(51) 国际专利分类号:

H04L 41/16 (2022.01)

(21) 国际申请号:

PCT/CN2023/123047

(22) 国际申请日:

2023 年 9 月 29 日 (29.09.2023)

(25) 申请语言:

中文

(26) 公布语言:

中文

(71) 申请人: 深圳 **TCL** 新技术有限公司
(**SHENZHEN TCL NEW TECHNOLOGY CO., LTD.**)
[CN/CN]; 中国广东省深圳市南山区西丽街道中山园路1001号国际E城D4栋9楼 518052 (CN)。

(72) 发明人: 曲淼(**QU, Miao**); 中国广东省深圳市南山区西丽街道中山园路1001号国际E城D4栋9楼 518052 (CN)。 陈喆(**CHEN, Zhe**); 中国广东省深圳市南山区西丽街道中山园路1001号国际E城D4

栋9楼 518052 (CN)。 张银成(**ZHANG, Yincheng**); 中国广东省深圳市南山区西丽街道中山园路1001号国际E城D4栋9楼 518052 (CN)。

(74) 代理人: 深圳紫藤知识产权代理有限公司(**PURPLEVINE INTELLECTUAL PROPERTY (SHENZHEN) CO., LTD.**); 中国广东省深圳市南山区粤海街道大冲社区华润置地大厦C座2901 518052 (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,

(54) **Title:** CONTROL METHOD FOR AI/ML MODEL IN WIRELESS COMMUNICATION SYSTEM

(54) 发明名称: 无线通信系统中的AI/ML模型的控制方法

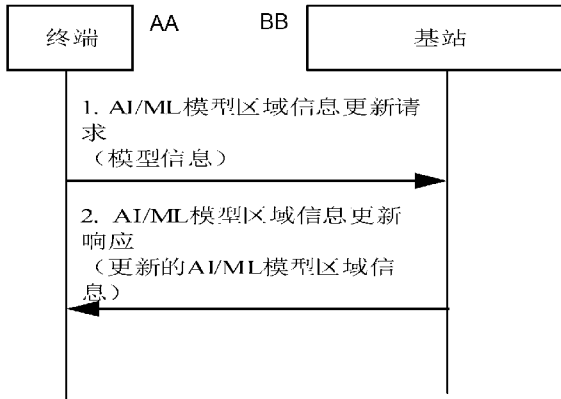


图 1

- AA Terminal
BB Base station
1 AI/ML model region information updating request (model information)
2 AI/ML model region information updating response (updated AI/ML model region information)

(57) **Abstract:** Provided in the present disclosure is a control method for an artificial intelligence/machine learning (AI/ML) model in a wireless communication system, which comprises: a node in the wireless communication system acquires AI/ML model region information, the node comprising at least one of a terminal or a base station; and the node determines model control for the AI/ML model on the basis of the AI/ML model region information. Further provided in the present disclosure is a wireless communication method for a terminal, which comprises: the terminal acquires from a first base station the AI/ML model region information, the AI/ML model region information being used for the terminal to determine the model control for the AI/ML model, and the terminal determines, on the basis of the AI/ML model region information, whether to execute the model control for the AI/ML model.



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。

- (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))。

(57) 摘要: 本公开提供一种无线通信系统中的人工智能/机器学习AI/ML模型的控制方法, 包括: 由无线通信系统中的节点获取AI/ML模型区域信息, 节点包括终端或基站中的至少一者; 由节点基于AI/ML模型区域信息来确定AI/ML模型的模型控制。本公开还提供一种用于终端的无线通信方法, 包括: 所述终端从第一基站获取AI/ML模型区域信息, 所述AI/ML模型区域信息用于所述终端确定AI/ML模型的模型控制, 其中, 所述终端基于所述AI/ML模型区域信息确定是否执行所述AI/ML模型的模型控制。

说明书

无线通信系统中的 AI/ML 模型的控制方法

技术领域

5 本公开涉及通信系统领域，更具体地，涉及一种用于无线通信系统中的人工智能/机器学习（AI/ML）模型的控制方法及其芯片、计算机可读存储介质和计算机程序产品、用于用户装备（终端）的无线通信方法及其芯片、计算机可读存储介质和计算机程序产品、以及用于第一基站的无线通信方法及其芯片、计算机可读存储介质和计算机程序产品。

10 背景技术

我们的社会正在经历数字化革命，在过去几年中，人工智能（AI）和机器学习（ML）方法被广泛应用于各个行业，以推动创新和提高流程效率。然而，在网络设计方面，随着数据量和网络复杂性的急剧增加，传统方法在许多情况下将无法提供快速解决方案。因此，AI/ML 将成为推动未来无线通信网络性能改进的不可或缺的技术。

15 实际上，在通过空中接口使用 AI/ML 的无线通信系统中，由于用户装备（终端）移动性，信道可能变得更差，从而可能导致切换发生。当终端移动到新的服务基站（如 gNB）时，系统可以在某些用例中使用传统方法或基于 AI/ML 的方法。如果采用基于 AI/ML 的方法，为了确保服务的连续性，在切换过程中可能需要执行 AI/ML 模型的模型控制（例如，模型选择、切换、更新等等）。但是，目前的无线通信网络并不具备支持此类模型控制的能力。

20 需要注意，在此章节中描述的内容不构成或不被视为对任何现有技术的承认。

发明内容

为了克服上述技术问题中的至少一个，本公开提供了一种用于无线通信系统中的人工智能/机器学习（AI/ML）模型的控制方法，其能够确保在切换过程中模型控制正常工作、减少信令
25 开销，并且减少动态信息上报的延迟。

本公开的一个方面提供了一种无线通信系统中的人工智能/机器学习 AI/ML 模型的控制方法，包括：

所述无线通信节点获取 AI/ML 模型区域信息，无线通信节点基于所述 AI/ML 模型区域信

息确定 AI/ML 模型的模型控制，所述无线通信节点为基站或终端。

根据本公开的一个或多个实施例，所述 AI/ML 模型的模型控制包括以下各项中的至少一者：

AI/ML 模型激活、AI/ML 模型去激活、AI/ML 模型选择、AI/ML 模型切换、AI/ML 模型

5 回退、AI/ML 模型更新、AI/ML 模型训练、AI/ML 模型重训练、或 AI/ML 模型微调。

根据本公开的一个或多个实施例，所述 AI/ML 模型区域信息包括以下各项中的至少一者：

AI/ML 模型区域标识 ID，

位置信息，

场景信息，

10 无线通信节点能力信息，

和/或 AI/ML 模型信息。

根据本公开的一个或多个实施例，所述位置信息包括以下各项中的至少一者：

小区信息、AI/ML 模型区域标识 ID、终端位置信息、基站位置信息，场景信息、控制单元 CU ID、数据单元 DU ID。

15 根据本公开的一个或多个实施例，所述场景信息包括以下各项中的至少一者：

传输环境信息，速度信息，角度扩展信息，时延扩展信息，信噪比，信道接收功率，和/或干扰信息。

根据本公开的一个或多个实施例，所述无线通信节点能力信息包括以下各项中的至少一者：

计算能力，存储能力，AI/ML 模型支持能力，负载情况，和/或传输能力，AI/ML 模型支

20 持 AI/ML 使用场景和/或 AI/ML 模型支持 AI/ML 特征。详细而言，在一些示例中例如一些特定计算能力，一些特定存储能力，AI/ML 模型支持一些特定能力，一些特定负载情况，一些特定传输能力，AI/ML 模型支持一些特定使用场景和/或 支持支持一些特定特征。

无线网络能力节点能力信息；指示一个无线网络节点的能力，如节点支持的 AI/ML 模型的模型操作能力、节点的硬件能力信息（如存储能力，算力，电池能力等）、节点支持的 AI/ML
25 模型的功能性/特征，节点支持的 AI/ML 模型的数据，节点支持的 AI/ML 模型的场景等。

根据本公开的一个或多个实施例，所述 AI/ML 模型信息包括以下各项中的至少一者：

AI/ML 模型 ID 信息，和/或 AI/ML 模型元信息。

根据本公开的一个或多个实施例，所述 AI/ML 模型区域信息通过以下各项中的至少一者传输：

30 非接入层 NAS 消息、无线电资源控制 RRC 消息、介质接入控制控制元素 MAC-CE 消息、

下行链路控制信息 DCI 消息、上行链路信道、下行链路信道、系统消息、N2 消息和/或 Xn 消息

根据本公开的一个或多个实施例，由所述无线通信节点基于所述 AI/ML 模型区域信息确定 AI/ML 模型的模型控制包括以下各项中的至少一者：

5 如果所述无线通信节点的位置信息属于所述 AI/ML 模型区域信息中所指示的位置信息，则所述无线通信节点不执行 AI/ML 模型的模型控制；

如果所述无线通信节点的位置信息不属于所述 AI/ML 模型区域信息中所指示的位置信息，则所述无线通信节点执行 AI/ML 模型的模型控制；

10 如果所述无线通信节点的 AI/ML 模型信息属于所述 AI/ML 模型区域信息中所指示的 AI/ML 模型信息，则所述无线通信节点不执行 AI/ML 模型的模型控制；

如果所述无线通信节点的 AI/ML 模型信息不属于所述 AI/ML 模型区域信息中所指示的 AI/ML 模型信息，则所述无线通信节点执行 AI/ML 模型的模型控制。

根据本公开的一个或多个实施例，所述执行 AI/ML 模型的模型控制包括以下各项中的至少一者：

15 所述执行 AI/ML 模型的模型控制包括以下各项中的至少一者：

AI/ML 模型激活、AI/ML 模型去激活、AI/ML 模型选择、AI/ML 模型切换、AI/ML 模型回退、AI/ML 模型更新、AI/ML 模型训练、AI/ML 模型重训练、或 AI/ML 模型微调、AI/ML 模型的模型控制的结果上报，和/或 AI/ML 模型的模型控制的响应。

20 根据本公开的一个或多个实施例，当所述无线通信节点为终端时，所述终端获取所述 AI/ML 模型区域配置信息，所述终端根据所述 AI/ML 模型区域信息执行所述 AI/ML 模型的模型控制。

根据本公开的一个或多个实施例，所述终端获取所述 AI/ML 模型区域配置信息包括以下各项中的至少一者：

AI/ML 模型信息、事由、指示符、终端 ID、和/或位置信息。

25 根据本公开的一个或多个实施例，当所述无线通信节点为基站时，所述基站获取所述 AI/ML 模型区域配置信息，所述基站根据所述 AI/ML 模型区域信息执行所述 AI/ML 模型的模型控制。

根据本公开的一个或多个实施例，所述基站获取所述 AI/ML 模型区域配置信息包括以下各项中的至少一者：

30 AI/ML 模型信息、事由、指示符、基站 ID，和/或位置信息。

根据本公开的一个或多个实施例，第一无线通信节点向第二无线通信节点发送的所述 AI/ML 模型区域信息请求消息和/或响应消息包括以下各项中的至少一者：

AI/ML 模型区域标识 ID，

位置信息，

5 场景信息，

无线通信节点能力信息，和/或

AI/ML 模型信息。

根据本公开的一个或多个实施例，所述无线通信节点发送所述 AI/ML 模型区域信息的请求消息，包括周期性发送或事件触发。

10 根据本公开的一个或多个实施例，在所述周期性发送的情况下，所述 AI/ML 模型区域信息响应于周期性配置而进行更新，其中，所述周期性配置包括所述更新的持续时间、起始时间、和/或结束时间。

根据本公开的一个或多个实施例，在所述方式为事件触发的情况下，所述 AI/ML 模型区域信息响应于事件触发配置而进行更新。

15 根据本公开的一个或多个实施例，所述周期性发送和/或所述事件触发相应的周期性配置和/或事件触发配置是预配置的、和/或配置的、和/或固定的。

根据本公开的一个或多个实施例，所述 AI/ML 模型区域信息请求消息和/或响应消息通过以下至少一项承载：

非接入层 NAS 消息、无线电资源控制 RRC 消息、介质接入控制控制元素 MAC-CE 消息、

20 下行链路控制信息 DCI 消息、上行链路信道、下行链路信道、系统消息、N2 消息和/或 Xn 消息。

本公开的一个方面提供了一种用于终端的无线通信方法，包括：

所述终端从第一基站获取 AI/ML 模型区域信息，所述 AI/ML 模型区域信息用于所述终端确定 AI/ML 模型的模型控制，

25 其中，所述终端基于所述 AI/ML 模型区域信息确定是否执行所述 AI/ML 模型的模型控制。

根据本公开的一个或多个实施例，还包括：

所述终端从所述第一基站获取 AI/ML 模型区域配置消息。

根据本公开的一个或多个实施例，如果第二基站不在所述 AI/ML 模型区域的情况下，所述终端确定执行所述 AI/ML 模型的模型控制。

30 根据本公开的一个或多个实施例，所述执行 AI/ML 模型的模型控制包括在所述终端与所

述第一基站之间更新所述 AI/ML 模型区域信息。

根据本公开的一个或多个实施例，还包括：

所述终端将针对所述 AI/ML 模型区域配置消息的上报消息以及所述执行所述 AI/ML 模型的模型控制的结果发送至所述第一基站，所述第一基站将获取的执行所述 AI/ML 模型的模型控制的结果发送至所述第二基站。

本公开的一个方面提供了一种用于第一基站的无线通信方法，包括：

所述第一基站从所述终端获取 AI/ML 模型的模型控制上报消息，所述 AI/ML 模型的模型控制上报消息包括所述 AI/ML 模型的模型控制的结果；

所述第一基站将所述第二基站的小区切换请求以及所述 AI/ML 模型的模型控制的结果发送至所述第二基站。

本公开的一个方面提供了一种用于第一基站的无线通信方法，包括：

所述第一基站根据终端反馈的 AI/ML 模型的模型控制消息，确定第二基站是否在 AI/ML 模型区域信息中；

如果所述第二基站不在所述 AI/ML 模型区域信息中，所述第一基站执行 AI/ML 模型的模型控制。

根据本公开的一个或多个实施例，所述 AI/ML 模型的模型控制包括在所述第一基站与所述第二基站之间更新所述 AI/ML 模型区域信息。

根据本公开的一个或多个实施例，还包括：

所述第一基站将所述第二基站的小区切换请求以及所述模型控制的结果发送至所述第二基站。

根据本公开的一个或多个实施例，还包括：

如果所述第二基站在所述 AI/ML 模型区域信息中，所述第一基站将所述 AI/ML 模型区域信息中所包含的 AI/ML 模型信息以及至所述第二基站的小区切换请求发送至所述第二基站。

本公开的一个方面提供了一种芯片，包括：

处理器，其被配置为调用和运行存储在存储器中的计算机程序，以使得其中安装有所述芯片的设备执行示例实施例中任一个的方法。

本公开的一个方面提供了一种芯片，包括：处理器，其被配置为调用和运行存储在存储器中的计算机程序，以使得其中安装有所述芯片的设备执行本公开任一方面的实施例的方法。

本公开的一个方面提供了一种计算机可读存储介质，其中存储有计算机程序，其中所述计算机程序使得计算机执行本公开任一方面的实施例的方法。

本公开的一个方面提供了一种计算机程序产品，包括计算机程序，其中所述计算机程序使得计算机执行本公开任一方面的实施例的方法。

附图说明

5 为了更清楚地说明本公开或相关技术的实施例，将在实施例中简要介绍以下附图。显然，附图仅仅是本公开的一些实施例，本领域的普通技术人员可以在不付出创造性劳动的前提下根据这些附图获得其他附图。

图 1 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

图 2 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

10 图 3 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

图 4 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

图 5 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

图 6 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

图 7 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

15 图 8 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

图 9 图示了根据本公开的一些实施例的在切换过程期间的快速模型控制过程的信令图。

图 10 图示了根据本公开的一些实施例的在切换过程期间的快速模型控制过程的信令图。

图 11 图示了根据本公开的一些实施例的在切换过程期间的快速模型控制过程的信令图。

图 12 是根据本公开的实施例的用于无线通信的示例系统的框图。

20

具体实施方式

本公开的实施例参照所述附图详细描述了技术事项、结构特征、实现的目的和效果，如下文所描述。具体地，本公开的实施例中的术语仅用于描述特定实施例的目的，而不是限制本公开。

25 在本公开中，“A 或 B”可以意指“仅 A”、“仅 B”或“A 和 B 二者”。

换句话说，在本公开中，“A 或 B”可以被解释为“A 和/或 B”。例如，在本公开中，“A、B 或 C”可以意指“仅 A”、“仅 B”、“仅 C”或“A、B、C 的任何组合”。

在本公开中使用的斜杠(/)或逗号可以意指“和/或”。例如，“A/B”可以意指“A 和/或 B”。因此，“A/B”可以意指“仅 A”、“仅 B”或“A 和 B 二者”。例如，“A、B、C”可以意指“A、B 或 C”。

在本公开中，“A 和 B 中的至少一个”可以意指“仅 A”、“仅 B”或“A 和 B 二者”。另外，在本公开中，表述“A 或 B 中的至少一个”或“A 和/或 B 中的至少一个”可以被解释为“A 和 B 中的至少一个”。

另外，在本公开中，“A、B 和 C 中的至少一个”可以意指“仅 A”、“仅 B”、“仅 C”或“A、B 和 C 的任何组合”。另外，“A、B 或 C 中的至少一个”或“A、B 和/或 C 中的至少一个”可以意指“A、B 和 C 中的至少一个”。

此外，术语“第一”、“第二”仅用于描述目的，而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或者隐含地包括一个或者多个所述特征。在本申请的描述中，“多个”的含义是两个或两个以上，除非另有明确具体的限定。

在通过空中接口使用 AI/ML 的无线通信系统中，由于终端移动性，信道可能变得更差，从而可能导致切换发生。当终端移动到新的服务基站（如 gNB）时，系统可以在某些用例中使用传统方法或基于 AI/ML 的方法。如果采用基于 AI/ML 的方法，为了确保服务的连续性，在切换过程中可能需要执行 AI/ML 模型的模型控制（例如，AI/ML 模型激活、AI/ML 模型去激活、AI/ML 模型选择、AI/ML 模型切换、AI/ML 模型回退、AI/ML 模型更新、AI/ML 模型训练、AI/ML 模型重训练、或 AI/ML 模型微调等等，本公开对此不进行任何限制）。但是，目前的无线通信网络并不具备支持此类模型控制的能力。

根据协议的进展，判断终端侧的 AI/ML 模型是否需要进行模型控制（也可称为模型操作、模型管理等）主要基于模型监听的结果。也就是说，在 AI/ML 模型的生命周期管理过程中，诸如模型激活、去激活、选择、切换、回退、更新、训练、重训练等 AI/ML 模型的模型控制是在模型开启监听后根据监听结果做出的决策。在切换过程中，如果采用模型监听的方案来判断当前的 AI/ML 模型是否适配于第二基站，需要第二基站配置模型监听和监听上报。然而，这个配置过程通常需要在终端接入第二基站后，也就是切换完成后才能进行模型监听的判断。而且，要得到准确的模型监听结果，需要通过多次基于 AI 的测量和传统测量来进行模型监听。

此外，在不同的场景下，对于终端侧的模型，当模型监听位于终端侧，监听决策位于网络（例如但不限于基站（如 gNB）等）侧时，终端需要上报监听结果以辅助网络侧进行新模型的选择，网络侧然后将新选择的模型下发给终端。当模型监听位于终端侧，监听决策也位于终端侧时，终端需要根据决策重新选择模型。

因此，对于切换来说，如果采用模型监听的方式，需要在终端接入新的第二基站后才能进行，并且整个模型监听到模型控制的过程会占用很长时间，给整个系统带来较大的延迟，无法

满足对服务连续性的要求。

另一种不采用模型监听的可选方案，即在发生小区切换时，为了使终端的 AI/ML 模型适配新的网络侧（如第二基站），可以直接进行 AI 模型的模型控制。例如新切换/选择/更新的 AI 模型可以由第二基站选择后直接传输给终端，或者由第二基站配置新模型的相关信息（如模型标识符（ID）、模型的元信息等）发送给终端，以辅助终端侧选择/切换新的 AI 模型。然而，这种方法会导致基站和终端之间需要直接交互 AI 模型或所需的 AI 模型信息，从而产生较大的信令开销。此外，考虑到新模型的传输和配置也是在终端接入新的第二基站后进行的，整个系统会面临较大的延迟，无法满足对服务连续性的要求。

根据 RAN1 和 RAN2 的进展，为了降低由于 AI 模型控制引起的切换时延问题，一些公司提出了在切换准备过程中交互模型信息的方法。通过这种方法，在切换可能发生时，可以使终端在切换后快速切换到正确的新的 AI/ML 上下文。虽然这种方法可以有效降低时延，但是执行模型控制的条件并不清楚。此外，在准备阶段存在大量的交互，如果当前的 AI/ML 模型仍然适用，那么会造成信令开销的浪费。

基于上述讨论，发明人洞察到设计一种更简单的方法来实现快速的模型控制是亟需且有实际意义的。

因而，本公开提出了一种用于无线通信系统中的人工智能/机器学习（AI/ML）模型的控制方法。该方法可以用于确定在切换过程中是否需要执行模型控制，并且可以用于在发生模型控制时辅助节点例如快速选择/切换/更新一个或多个合适的 AI/ML 模型。借助于本公开的方法，能够有效降低由于模型监听以及对 AI/ML 模型进行模型控制等所带来的延迟；并且能够降低执行模型控制时的信令交互开销。

在一些实施例中，终端可以为用户站、移动站、移动台、无线通信设备、用户代理等各种用户设备。此外，终端的一些示例可以为手机、平板电脑、笔记本电脑、掌上电脑、移动互联网设备、可穿戴设备、虚拟现实设备、增强现实设备等，并且本公开对此并不进行任何限定。

在一些实施例中，基站可以包括或可以被称作基站收发机、无线基站、接入点、无线收发机、节点 B、e 节点 B（eNB）、下一代节点 B 或千兆节点 B（其中任一个可以被称作 gNB）、家庭节点 B、家庭 e 节点 B 或其他合适的术语。无线通信系统可以包括不同类型的基站（例如，宏小区基站或小型小区基站等）。终端可以与各种基站和/或网络设备通信，包括宏 eNB、小型小区 eNB、gNB、中继基站等。

提供了一种 AI/ML 模型特定的区域信息，其可用于辅助节点（如终端、gNB 等）执行 AI/ML 模型控制。其中，一个 AI/ML 模型可以用于特定 AI/ML 区域中的一个或多个位置。在一些实

施例中，AI/ML 模型区域信息可以至少包括以下内容之一：AI/ML 模型区域标识 ID，位置信息，场景信息，无线通信节点能力信息，和/或 AI/ML 模型信息。在一些实施例中，AI/ML 模型区域信息可以通过 RRC 消息、NAS 消息、N2 消息、Xn 消息、DCI、MAC CE，系统消息，上行链路信道、下行链路信道等配置给节点，如终端和/或基站等。

5

部分本方面实施例出现的术语解释如下：

模型控制：也可称为模型操作、模型管理等，指的是生命周期管理（LCM）中的一些模型操作，例如模型选择、模型切换、模型更新、模型激活、模型去激活（模型停用）、模型回退、模型保留（微调）、模型训练、模型重训练等，本公开对此不进行任何限制；

10

终端侧（端）模型：可以指位于终端侧（端）的 AI/ML 模型，包括终端侧模型之情形和双侧模型中的终端部分之情形；

AI/ML 模型区域信息：是指 AI 特定区域（AI-specific area），在本公开中，也可以被可互换地指称为 AI/ML 模型特定区域，AI/ML 模型区域信息与 AI/ML 模型相对应 或相关联；

15

AI/ML 模型区域指示符：用于标识 AI 特定区域；

AI/ML 模型特定区域的特定区域代码（AI-specific area code of AI-specific area）：可以用于指示 AI 特定区域的 AI 特定区域代码；

AI/ML 模型信息：可以用于指示 AI/ML 模型的信息，例如模型 ID、模型元信息等，此外，当没有特别指出时，AI/ML 模型信息可以是一个 AI/ML 模型或多个 AI/ML 模型的信息；

20

位置信息：小区信息、AI/ML 模型区域标识 ID、终端位置信息、基站位置信息，场景信息、控制单元 CU ID、数据单元 DU ID 等；其中，

根据本公开的一个或多个实施例，所述无线通信节点能力信息包括以下各项中的至少一者：

25

计算能力，存储能力，AI/ML 模型支持能力，负载情况，和/或传输能力，AI/ML 模型支持 AI/ML 使用场景和/或 AI/ML 模型支持 AI/ML 特征。详细而言，在一些示例中例如一些特定计算能力，一些特定存储能力，AI/ML 模型支持一些特定能力，一些特定负载情况，一些特定传输能力，AI/ML 模型支持一些特定使用场景和/或 支持支持一些特定特征。

30

无线网络能力节点能力信息；指示一个无线网络节点的能力，如节点支持的 AI/ML 模型的模型操作能力、节点的硬件能力信息（如存储能力，算力，电池能力等）、节点支持的

AI/ML 模型的功能性/特征，节点支持的 AI/ML 模型的数据，节点支持的 AI/ML 模型的场景等。

基站信息：可以指示小区 ID、传输接收点（TRP）ID、中央单元（CU）ID 和/或分布单元（DU）ID 等；

5 场景信息：可以传输环境信息，速度信息，角度扩展信息，时延扩展信息，信噪比，信道接收功率，和/或干扰信息等反映信道特征的一个或多个参数；

终端位置信息：可以包括例如终端 ID、与终端相关联的小区 ID、终端的物理位置区域、经度、纬度、海拔高度等。

模型 ID：用于标识 AI/ML 模型，包括全局模型 ID 和/或本地模型 ID，其中：

10 全局 AI/ML 模型 ID 用于标识 AI/ML 模型，其可以是全球唯一 ID、PLMN 特定唯一 ID、运营商特定唯一 ID 或 AI/ML 管理平台特定唯一 ID。

本地 AI/ML 模型 ID 用于标识网络中使用的 AI/ML 模型，并与全局 ID 存在一定的映射关系。

15 逻辑 ID 可以是全球唯一 ID、PLMN 特定唯一 ID、运营商特定唯一 ID、小区特定 ID、链路特定 ID、跟踪区域（TA）特定 ID、CU 特定 ID、DU 特定 ID、用户面功能（UPF）特定 ID、接入和移动性管理功能（AMF）特定 ID、无线电资源控制（RRC）特定 ID 或网络切片特定 ID。

20 术语“小区特定”表示每个 AI/ML 模型在特定小区内具有唯一 ID。类似地，“链路特定”、“TA 特定”、“CU 特定”、“DU 特定”、“UPF 特定”、“AMF 特定”和“网络切片特定”等指的是网络中特定上下文中 AI/ML 模型的唯一 ID。

模型 ID 也可以通过隐式方式指示，例如 RRC ID、DCI ID 等。

模型元信息：用于描述 AI/ML 模型的信息，可以包括模型功能性/特征、模型版本信息、模型格式信息、供应商信息、计算复杂度、模型复杂度、模型性能、适用条件等。其中，

25 模型功能性/特征：指示一个模型的功能性，例如模型用于 CSI 压缩、波束预测、定位等。指示应该是子用例级别或甚至更细的级别。

模型版本信息：指示 AI/ML 模型的版本信息。

模型格式信息：指示 AI/ML 模型的格式信息，例如开放格式（如 ONNX）、私有格式等。

供应商信息：指示 AI/ML 模型训练供应商的信息。

30 计算复杂度：用于指示模型训练的计算复杂度，例如 FLOP（浮点运算数）、预处理

理/后处理的级别。

模型复杂度：用于指示模型训练的复杂度，包括实值模型参数的数量、实值的数量等。

模型性能：用于指示模型的性能，其可以是模型准确度、模型偏差、模型方差等。

5 适用条件：用于指示 AI/ML 模型或功能性的适用条件，例如场景、配置或站点、终端内部能力等。

场景信息包括以下各项中的至少一者：传输环境信息，速度信息，角度扩展信息，时延扩展信息，信噪比，信道接收功率，和/或干扰信息，其中，

10 传输环境信息：用于描述当前无线通信节点所处的信道传输环境，比如处于室内/城市/城市郊区/农村/隧道等，或者信道分类信息，比如基于 UMA/UM i/InH/RMA 的分类信息等。

速度信息：用于描述当前无线通信节点所处的移动速度，该移动速度可以是终端测量得到的，也可以是基站测量得到的。

15 角度扩展信息：用于描述当前无线通信节点所出的信道的角度扩展情况，即信道的角度扩展实时信息或者统计信息，或者基于角度扩展相关联的其他信息。

时延扩展信息：用于描述当前无线通信节点所出的信道的时延扩展情况，即信道的时延扩展实时信息或者统计信息，或者基于时延扩展相关联的其他信息。

20 根据本公开的一个或多个实施例，所述无线通信节点能力信息包括以下各项中的至少一者：

计算能力，存储能力，AI/ML 模型支持能力，负载情况，和/或传输能力，AI/ML 模型支持 AI/ML 使用场景和/或 AI/ML 模型支持 AI/ML 特征。详细而言，在一些示例中例如一些特定计算能力，一些特定存储能力，AI/ML 模型支持一些特定能力，一些特定负载情况，一些特定传输能力，AI/ML 模型支持一些特定使用场景和/或 支持支持一些特定特征。

25 无线网络能力节点能力信息：指示一个无线网络节点的能力，如节点支持的 AI/ML 模型的模型操作能力、节点的硬件能力信息（如存储能力，算力，电池能力等）、节点支持的 AI/ML 模型的功能性/特征，节点支持的 AI/ML 模型的数据，节点支持的 AI/ML 模型的场景等。

根据本公开的一个或多个实施例，所述 AI/ML 模型信息包括以下各项中的至少一者：

AI/ML 模型 ID 信息，和/或 AI/ML 模型元信息。

30

如本公开所使用的，术语“模型控制”也可称为模型操作、模型管理等。

如本公开所使用的，当提到表述“不发起”时，可以指禁止执行、不执行、放弃执行等。

AI/ML 模型区域信息配置

5 AI 特定区域信息被引入用于辅助节点（如终端、记账等）正确执行 AI/ML 模型控制，和/或在模型控制发生时帮助节点例如选择/切换/更新一个或多个合适的 AI/ML 模型。

关于本公开所提出的 AI/ML 模型区域信息，以下问题和相关方法列举如下：

AI/ML 模型区域信息可以由以下的方式构成。

10 AI/ML 模型区域信息的组成包括以下各项中的至少一者：AI/ML 模型区域标识 ID，位置信息，场景信息，无线通信节点能力信息，和/或 AI/ML 模型信息。几种可能的选项如下

- 选项 1：AI/ML 模型区域信息可以由单个或多个位置信息构成；
- 选项 2：AI/ML 模型区域信息可以由单个或多个位置信息和一些 AI/ML 模型信息构成；
- 选项 3：AI/ML 模型区域信息可以由单个或多个位置信息，和一些 AI/ML 模型信息，和一些场景信息构成。

15 AI/ML 模型区域标识 ID，位置信息，场景信息，无线通信节点能力信息，和/或 AI/ML 模型信息任意两项组合及以上存在映射关系，所述映射关系可以预先配置或通过例如 RRC 消息、NAS 消息、DCI 消息等进行配置，也可以在特定情况下被固定。在一种实现方式种，所述 AI/ML 模型信息包括以下各项中的至少一者：AI/ML 模型 ID 信息，和/或 AI/ML 模型元信息。

20 AI/ML 模型区域信息与 AI/ML 模型信息之间的映射可以是一对一、一对多或多对一的关系。需要注意，AI/ML 模型区域信息中的任意一个或多个元素与 AI/ML 模型信息中的任意一个或多个元素之间都可以存在映射关系，本公开对此不进行任何限制。

25 在一些示例中，可以使用 AI/ML 模型区域指示符表示位置信息和 AI/ML 模型信息之间的映射。表 1 展示了 AI/ML 模型区域信息与 AI/ML 模型信息之间的一种可能的映射关系，其中位置信息是小区 ID，模型信息是模型 ID；表 2 展示了 AI/ML 模型区域信息与 AI/ML 模型信息之间的另一种可能的映射关系，其中位置信息是小区 ID，模型信息是模型特征和/或功能性、模型格式。

表 1：AI/ML 模型区域信息与 AI/ML 模型信息之间的映射关系

AI/ML 模型区域指示符	位置信息	AI/ML 模型信息
1	小区 ID#1, 小区#2, 小区#3	模型#1, 模型#2, 模型#3,

2	小区 ID#1, 小区#4	模型#4,
3	小区 ID#1, 小区#4	模型#1, 模型#5

表 2: AI/ML 模型区域信息与 AI/ML 模型信息之间的映射关系

AI/ML 模型区域指示符	位置信息	AI/ML 模型信息	
1	小区 ID#1, 小区#2, 小区#3	模型特征: CSI 压缩	模型格式: 开放格式 (ONNX)
2	小区 ID#1, 小区#4	模型特征: BM 预测	模型格式: 专有格式
3	小区 ID#1, 小区#4	模型特征: BM 预测	模型格式: 专有格式

5 在一些示例中, AI/ML 模型区域信息与 AI/ML 模型信息之间的一种可能映射关系如下表 3 所示, 其中位置信息是小区 ID, 模型信息是模型 ID:

表 3: AI/ML 模型区域信息与 AI/ML 模型信息之间的映射关系

AI/ML 模型区域指示符	AI/ML 模型信息
1 (小区 ID#1, 小区#2, 小区#3)	模型#1, 模型#2, 模型#3,
2 (小区 ID#1, 小区#4)	模型#4,
3 (小区 ID#1, 小区#4)	模型#1, 模型#5

10 在一些示例中, AI/ML 模型区域信息与 AI/ML 模型信息之间的一种可能映射关系如下表 4 所示, 其中位置信息是小区 ID, 模型信息是模型 ID 和 AI 特征:

表 4: AI/ML 模型区域信息与 AI/ML 模型信息之间的映射关系

AI/ML 模型区域指示符	AI/ML 模型信息
1 (小区 ID#1, 小区#2, 小区#3)	模型#1, 模型#2, 模型#3, 波束管理预测
2 (小区 ID#1, 小区#4)	模型#4, 波束管理预测
3 (小区 ID#1, 小区#4)	模型#1, 模型#5, CSI 预测

在一些示例中, 为了指示 AI/ML 模型区域信息与 AI/ML 模型信息之间的一种可能映射关系, 可以引入一个索引, 如下表 5 所示, 其中位置信息是小区 ID, 模型信息是模型 ID:

表 5: AI/ML 模型区域信息与 AI/ML 模型信息之间的映射关系

索引	AI/ML 模型区域指示符	AI/ML 模型信息
----	---------------	------------

1	1 (小区 ID#1, 小区#2, 小区#3)	模型#1, 模型#2, 模型#3,
2	2 (小区 ID#1, 小区#4)	模型#4

应当领会，上面描述的其中构成 AI/ML 模型区域信息的几种不同方式的示例仅出于说明性目的而被给出，本公开对此并不旨在进行任何限制。本领域技术人员在阅读上述示例后，容易想到其他合适的添加、删除、修改、变型等，而这些都落入本公开的范围。

5 关于 AI/ML 模型区域信息的配置，有若干种不同的选择：

在 RRC 激活、非激活或空闲状态下的终端（如 UE）可以配置有 AI/ML 模型区域信息或 AI/ML 模型区域信息列表，可以通过以下选项进行配置：

选项 1：核心网网元（CN，如 AMF）通过例如 NAS 消息向终端发送 AI/ML 模型区域信息或 AI/ML 模型区域信息列表。本公开对所使用的消息类型不进行任何限制，并且不排除使用除 NAS 消息之外的任何其他合适的消息；

选项 2：基站（如 gNB）通过例如 RRC 消息向终端发送 AI/ML 模型区域信息或 AI/ML 模型区域信息列表，RRC 消息例如包括 RRC 重配置消息、RRC 释放消息或其他专用的 RRC 消息。本公开对所使用的消息类型不进行任何限制，并且不排除使用除 RRC 消息之外的任何其他合适的消息；

15 选项 3：基站（如 gNB）通过系统信息消息向终端发送 AI/ML 模型区域信息或 AI/ML 模型区域信息列表，包括 SIB1 或其他新定义的 SIB。本公开对所使用的消息类型不进行任何限制，并且不排除使用除系统信息消息之外的任何其他合适的消息。

选项 4：基站（如 gNB）通过例如 MAC CE 消息向终端发送 AI/ML 模型区域信息或 AI/ML 模型区域信息列表；本公开对所使用的消息类型不进行任何限制，并且不排除使用除 MAC CE 消息之外的任何其他合适的消息；

20 选项 5：基站（如 gNB）通过例如 DCI 消息向终端发送 AI/ML 模型区域信息或 AI/ML 模型区域信息列表；本公开对所使用的消息类型不进行任何限制，并且不排除使用除 DCI 消息之外的任何其他合适的消息；

25 一个基站（如 gNB）可以配置有 AI/ML 模型区域信息或 AI/ML 模型区域信息列表，可以通过以下选项进行配置：

选项 1：核心网网元（如 AMF）通过 N2 消息向基站（如 gNB）发送 AI/ML 模型区域信息或 AI/ML 模型区域信息列表。本公开对所使用的消息类型不进行任何限制，并且不排除使

用除 N2 消息之外的任何其他合适的消息；

选项 2：基站（如 gNB）通过 Xn 消息向另一个基站（如 gNB）发送 AI/ML 模型区域信息或 AI/ML 模型区域信息列表。本公开对所使用的消息类型不进行任何限制，并且不排除使用除 Xn 消息之外的任何其他合适的消息。

5 应当领会，上面描述的其中配置 AI/ML 模型区域信息的几种不同方式的示例仅出于说明性目的而被给出，本公开对此并不旨在进行任何限制。本领域技术人员在阅读上述示例后，容易想到其他合适的添加、删除、修改、变型等，而这些都落入本公开的范围。

关于 AI/ML 模型区域信息更新，可以考虑至少三种替代方案：

替代方案 1：周期性请求更新，可以通过例如 RRC 消息、NAS 消息、N2 消息或系统消息，
10 Xn 消息，MAC CE，DCI 等配置 AI/ML 模型区域信息更新周期，更新周期配置内容包括以下至少一个更新周期的持续时间、起始时间和结束时间，例如 5 分钟到 720 分钟。

替代方案 2：事件触发请求更新，以下任一事件可能触发 AI/ML 模型区域信息更新：

事件 1：AI/ML 模型区域信息对应的模型不在配置的 AI/ML 模型区域信息中，将触发 AI/ML 模型区域信息更新；例如，如果模型发生变化（包括模型选择、模型切换、模型更新等），对
15 应的 AI/ML 模型区域信息不属于配置的 AI/ML 模型区域信息。

事件 2：候选基站和/或小区的位置信息不在配置的 AI/ML 模型区域信息中，将触发 AI/ML 模型区域信息更新。

事件 3：当终端的小区重选过程选择了不属于配置的 AI/ML 模型区域信息的基站和/或小区，将触发 AI/ML 模型区域信息更新。

20 下面描述 AI/ML 模型区域信息更新的可能过程。

图 1 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

如图 1 所示，图示了终端触发的 AI/ML 模型区域信息更新过程。该过程可以由事件触发或周期触发。

根据一些实施例，提供了一种用于更新 AI/ML 模型区域信息的方法。以下是该方法的步骤/信令：
25

示例性的，步骤 1 可以包括终端（如终端）向基站（如 gNB）发送用于 AI/ML 模型区域信息更新的请求消息。该消息至少包括以下信息：

- 模型信息：用于指示终端侧当前的 AI 模型，可以是处于激活态的 AI 模型，也可以处于非激活态的 AI 模型。
- 30 - 事由：用于指示 AI/ML 模型区域信息更新的事由。在本例中，事由是终端侧当前 AI

模型与 AI/ML 模型区域信息不属于配置的 AI/ML 模型区域信息列表。

- 终端标识：用于标识特定的终端。

示例性的，该消息可以是专用的用于请求 AI/ML 模型区域信息更新的消息。一旦基站（如 gNB）接收到该消息，基站（如 gNB）就知道该消息是用于请求 AI/ML 模型区域信息更新的。

- 5 或者，该消息可以是一个传统消息，其可包含一个指示符，用于指示该消息用于 AI/ML 模型区域信息更新请求。另外，该消息也可以是一个传统消息，其可包含一个事由，用于指示发起专用消息的原因。

示例性的，步骤 2 可以包括终端接收到更新的 AI/ML 模型区域信息。更具体地说，更新的 AI/ML 模型区域信息可以包括 AI/ML 模型区域指示符、AI/ML 模型区域指示符列表、AI 模型信息、位置信息或 AI/ML 模型区域信息与模型信息的映射规则/关系。通过接收这些信息，
10 终端能够更新其 AI/ML 模型区域信息，并获得最新的 AI 模型信息和相关的位置信息和映射规则。这样，终端能够更好地适应当前的网络环境和需求，提供更高质量的服务。因此，本公开的实施例提供了一种有效的 AI/ML 模型区域信息更新方法，能够提升网络性能和用户体验。

需要注意，上述实施例仅为示例，并不限制本发明的范围。在不脱离本发明精神和范围的情况下，
15 可以对实施例进行各种修改和变化。

图 2 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

如图 2 所示，图示了终端触发的 AI/ML 模型区域信息更新过程。该过程可以由事件触发或定期触发。

在某些情况下，如果第一基站（例如，在基站为 gNB 的情况下为服务基站 s-gNB）没有与请求的 AI/ML 模型区域信息对应的模型，则第一基站可以将请求转发给其他基站（例如，
20 第二基站），如图 2 所示。

根据一些实施例，提供了一种用于更新 AI/ML 模型区域信息的方法。以下是该方法的步骤/信令：

示例性的，步骤 1 可以包括终端向第一基站发送 AI/ML 模型区域信息更新的请求消息，
25 其信息与图 1 所示示例中的步骤 1 相同。

示例性的，步骤 2 可以包括第一基站将在步骤 1 中收到的请求消息转发给（例如多个）其他基站。其他基站会发送更新后的 AI/ML 模型区域信息，更具体地说，更新后的 AI/ML 模型区域信息可以包括 AI/ML 模型区域标识 ID，位置信息，场景信息，无线通信节点能力信息，和/或 AI/ML 模型信息。

30 示例性的，步骤 3 可以包括终端接收关于 AI/ML 模型区域信息更新请求的响应消息，其

中的信息与步骤 2 中其他基站发送给第一基站的信息相同。

通过上述示例过程，终端能够向基站，如第一基站以及进而其他基站请求 AI/ML 模型区域信息更新。通过接收更新后的 AI/ML 模型区域信息，终端可以确保具备最相关的 AI 模型，提高性能并适应不断变化的网络条件。因此，本公开的实施例提供了一种有效的终端触发的

5 AI/ML 模型区域信息更新过程，能够提升网络性能和用户体验。

需要注意，上述实施例仅为示例，并不限制本发明的范围。在不脱离本发明精神和范围的情况下，可以对实施例进行各种修改和变化。

图 3 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

如图 3 所示，图示了基站（如 gNB）触发的 AI/ML 模型区域信息更新过程。该过程可以

10 由事件触发或定期触发。

在某些情况下，如果第一基站（s-gNB）没有与其他候选 gNB 对应的 AI/ML 模型区域信息（或 AI/ML 模型区域信息列表），则第一基站会触发 AI/ML 模型区域信息更新过程。

根据一些实施例，提供了一种用于更新 AI/ML 模型区域信息的方法。以下是该方法的步骤/信令：

15 示例性的，步骤 1 可以包括第一基站触发 AI/ML 模型区域信息更新过程。具体来说，第一基站可以向其他基站发送 AI/ML 模型区域信息更新的请求消息，该信息可以包括模型信息和事由。

示例性的，步骤 2 可以包括其他基站向第一基站 B 发送关于 AI/ML 模型区域信息更新请求的响应消息。作为示例而非限制，响应消息可以指示 AI/ML 模型区域信息更新请求被准予，

20 并且包括更新后的 AI/ML 模型区域信息。

通过上述示例过程，基站（如 gNB）能够触发 AI/ML 模型区域信息的更新，以确保所配置的 AI/ML 模型区域信息（列表）的有效性，例如，包括与其他候选基站相对应的 AI/ML 模型区域信息（列表）。因此，本公开的实施例提供了一种由基站（如 gNB）触发的 AI/ML 模型区域信息更新过程，能够提升网络性能和效率。

25 需要注意，上述实施例仅为示例，并不限制本发明的范围。在不脱离本发明精神和范围的情况下，可以对实施例进行各种修改和变化。

图 4 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

如图 4 所示，图示了终端触发的 AI/ML 模型区域信息更新过程。该过程可以由事件触发或定期触发。

30 根据一些实施例，提供了一种用于更新 AI/ML 模型区域信息的方法。以下是该方法的步

骤/信令：

示例性的，步骤 1 可以包括终端向基站（如 gNB）发送 AI/ML 模型区域信息更新的请求消息，该消息至少包括以下内容：

位置信息：用于指示基站（如 gNB）的位置信息，包括候选基站。

- 5 事由：用于指示 AI/ML 模型区域信息更新的原因。在此情况下，事由可以是候选基站的小区不在所配置的 AI/ML 模型区域信息中，或者进行了模型（重新）选择，新的驻留的小区不属于所配置的 AI/ML 模型区域信息。

终端标识：用于指示终端的身份。

- 10 该消息可以是专用的用于请求 AI/ML 模型区域信息更新的消息。一旦基站（如 gNB）接收到该消息，就知道它是用于要求 AI/ML 模型区域信息更新的。另外，该消息也可以是传统消息，其可包含一个指示符，用于指示该消息用于 AI-TA 更新请求。或者，该消息也可以是传统消息，其可包含一个用于指示为什么要发起专用消息的原因。

- 15 示例性的，步骤 2 可以包括终端接收更新后的 AI/ML 模型区域信息，更具体地说，更新后的 AI/ML 模型区域信息可以包括以下任意之一：AI/ML 模型区域指示符、AI/ML 模型区域指示符列表、AI 模型信息、位置信息或 AI/ML 模型区域信息和模型信息的映射规则/关系。

通过上述示例过程，终端能够向基站（如 gNB）发送 AI/ML 模型区域信息更新请求，并接收更新后的 AI/ML 模型区域信息，以提高网络性能和效率。

需要注意，上述实施例仅为示例，并不限制本发明的范围。在不脱离本发明精神和范围的情况下，可以对实施例进行各种修改和变化。

- 20 图 5 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

如图 5 所示，图示了终端触发的 AI/ML 模型区域信息更新过程。该过程可以由事件触发或定期触发。

根据一些实施例，提供了一种用于更新 AI/ML 模型区域信息的方法。以下是该方法的步骤/信令：

- 25 示例性的，步骤 1 可以包括终端向基站（如 gNB）发送 AI/ML 模型区域信息更新的请求消息，该消息至少包括以下内容之一：位置信息、事由、终端标识等。更具体地说，对于上述事件 2 而言，候选基站的信息需要包含在消息中，以指示所请求的 AI/ML 模型区域信息（列表）。

- 30 示例性的，步骤 2 可以包括第一基站将步骤 1 中的请求消息转发给（例如多个）其他基站，其他基站发送更新后的 AI/ML 模型区域信息，更具体地说，更新后的 AI/ML 模型区域信息至

少包含以下之一：AI/ML 模型区域指示符、AI/ML 模型区域指示符列表、AI 模型信息、位置信息或 AI/ML 模型区域信息和模型信息的映射规则/关系。

示例性的，步骤 3 可以包括终端从第一基站接收更新后的 AI/ML 模型区域信息，更具体地说，更新后的 AI/ML 模型区域信息至少包含以下之一：AI/ML 模型区域指示符、AI/ML 模型区域指示符列表，AI 模型信息、位置信息或 AI/ML 模型区域信息和模型信息的映射规则/关系。

通过上述示例过程，终端能够向基站，如第一基站以及进而其他基站发送 AI/ML 模型区域信息更新请求，并从第一基站接收更新后的 AI/ML 模型区域信息，以提高网络性能和效率。

需要注意，上述步骤仅为示例，并不限制本发明的范围。在不脱离本发明精神和范围的情况下，可以对步骤进行各种修改和变化。

图 6 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

如图 6 所示，图示了基站（如 gNB）触发的 AI/ML 模型区域信息更新过程。该过程可以由事件触发或定期触发。

根据一些实施例，提供了一种用于更新 AI/ML 模型区域信息的方法。以下是该方法的步骤/信令：

示例性的，步骤 1 可以包括第一基站向其他基站发送 AI/ML 模型区域信息更新的请求消息，该消息至少包括以下内容之一：位置信息、事由、终端标识 等。更具体地说，对于事件 2 而言，候选基站的小区信息需要包含在消息中，以指示所请求的 AI/ML 模型区域信息（列表）。

示例性的，步骤 2 可以包括其他基站对步骤 1 中的消息进行响应，其他基站发送更新后的 AI/ML 模型区域信息，更具体地说，更新后的 AI/ML 模型区域信息可以是 AI/ML 模型区域指示符、AI/ML 模型区域指示符列表、AI 模型信息、位置信息或 AI/ML 模型区域信息和模型信息的映射规则/关系。

通过上述示例过程，第一基站能够向其他基站发送 AI/ML 模型区域信息更新请求，并接收来自其他基站的更新后的 AI/ML 模型区域信息，以提高网络性能和效率。

需要注意，上述步骤仅为示例，并不限制本发明的范围。在不脱离本发明精神和范围的情况下，可以对步骤进行各种修改和变化。

图 7 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

如图 7 所示，图示了终端触发的 AI/ML 模型区域信息更新过程。该过程可以由事件触发或定期触发。

在某些情况下，终端触发了 AI/ML 模型区域信息更新过程，它将向基站（如 gNB）发送

请求消息，基站（如 gNB）将消息转发给 CN 节点，CN 将通过更新 AI/ML 模型区域信息来响应终端。以下是一个关于小区重选的示例，AI/ML 模型区域信息更新可能会在注册过程中发生，AI/ML 模型区域信息由 AMF 发送给终端。

根据一些实施例，提供了一种用于更新 AI/ML 模型区域信息的方法。以下是该方法的步骤/信令：

示例性的，步骤 1 可以包括终端通过注册请求消息向基站（如 gNB）发送 AI/ML 模型区域信息更新的请求消息，该消息至少包括以下内容之一：位置信息、事由、终端标识、AI 模型信息等。更具体地说，对于事件 1 而言，AI 模型信息需要包含在消息中；对于事件 2 而言，候选基站的小区信息需要包含在消息中，以指示所请求的 AI/ML 模型区域信息（列表）。

示例性的，步骤 2 可以包括基站（如 gNB）通过例如 N2 消息将 AI/ML 模型区域信息更新请求消息转发给 AMF。

示例性的，步骤 3 可以包括如果终端的注册请求被接受，终端通过来自 AMF 的注册接受消息接收更新后的 AI/ML 模型区域信息。

通过上述示例过程，终端能够触发 AI/ML 模型区域信息更新过程，并从 AMF 接收更新后的 AI/ML 模型区域信息，以提高网络性能和效率。

需要注意，上述步骤仅为示例，并不限制本发明的范围。在不脱离本发明精神和范围的情况下，可以对步骤进行各种修改和变化。

图 8 图示了根据本公开的一些实施例的 AI/ML 模型区域信息更新过程的信令图。

如图 8 所示，图示了基站（如 gNB）触发的 AI/ML 模型区域信息更新过程。该过程可以由事件触发或定期触发。

在某些情况下，基站（如 gNB）触发了 AI/ML 模型区域信息更新过程，它将向 CN 节点（例如 AMF、网络数据分析功能（NWDAF）等）发送请求消息。

根据一些实施例，提供了一种用于更新 AI/ML 模型区域信息的方法。以下是该方法的步骤/信令：

示例性的，步骤 1 可以包括基站（如 gNB）通过例如 N2 消息向 AMF 发送 AI/ML 模型区域信息更新的请求消息。该请求消息至少包括以下内容之一：位置信息、事由、终端标识、AI 模型信息等。更具体地说，对于事件 1 而言，AI 模型信息需要包含在消息中；对于事件 2 而言，候选基站的小区信息需要包含在消息中，以指示所请求的 AI/ML 模型区域信息（列表）。

示例性的，步骤 2 可以包括基站通过例如 N2 消息从 AMF 接收更新后的 AI/ML 模型区域信息。

通过上述示例过程，基站（如 gNB）能够通过 N2 消息向 AMF 发送 AI/ML 模型区域信息更新的请求消息，并且能够通过 N2 消息从 AMF 接收更新后的 AI/ML 模型区域信息，以实现更高效的网络通信和优化。

需要注意，上述步骤仅为示例，并不限制本发明的范围。在不脱离本发明精神和范围的情况下，可以对步骤进行各种修改和变化。

终端在切换过程中确定快速模型控制的过程

图 9 图示了根据本公开的一些实施例的在切换过程期间的快速模型控制过程的信令图。

如图 9 所示，展示了潜在切换过程中的快速模型控制。对于本例，模型控制（例如模型选择、模型切换、模型更新等）位于终端端。

示例性的，步骤 0：终端接收有关 AI/ML 模型区域信息的配置消息，该消息用于在切换过程中帮助终端快速确定模型控制。该消息可包括 AI/ML 模型区域指示符、AI/ML 模型区域指示符列表、AI 模型信息、位置信息或 AI/ML 模型区域信息与模型信息的映射规则/关系，并且该消息可以在例如 RRC 消息、系统信息、MAC CE、DCI，NAS 等中传递。

示例性的，步骤 1：终端接收传统测量配置，测量配置通过专用信令（例如 RRCReconfiguration 或 RRCResume 等）提供。

示例性的，步骤 2：终端检查配置的 AI/ML 模型区域信息，如果候选第二基站（如 Target gNB,T-gNB）在配置的 AI/ML 模型区域信息列表范围内，则无需执行模型控制，然后终端（如终端）执行步骤 4；如果候选第二基站（如 Target gNB,T-gNB）不在配置的 AI/ML 模型区域信息列表范围内，则终端将执行步骤 3 模型控制，然后终端执行步骤 3'和步骤 4，步骤 3'的详细信息在上文描述。

示例性的，步骤 4：基于步骤 2 的检查结果，终端向第一基站发送测量上报消息，该消息可携带模型控制结果，消息中至少包括以下之一：模型信息、指示或原因。

指示：用于指示当前进行中的模型是否已更改。

模型信息：用于指示当前的 AI/ML 模型。

事由：用于指示 AI/ML 模型更改或未更改的原因。

示例性的，步骤 5：第一基站发起切换（Handover, HO）决策。

示例性的，步骤 6：第一基站向第二基站发送切换请求消息，该消息可携带模型信息，模型信息表示适用于第二基站的 AI/ML 模型，此步骤可以在终端和第二基站之间提前进行模型识别。

示例性的，步骤 7：第二基站执行接纳控制。

示例性的，步骤 8：第二基站向原基站发送关于切换请求的确认消息，其中携带第二基站的信息和终端的资源信息。

5 示例性的，步骤 9：第一基站向终端发送带有切换命令的消息（如 RRCReconfiguration 消息等），其中携带用于访问第二基站的信息，例如小区 ID。

需要注意，步骤 3'可以在步骤 4 之前或之后发生。在某些示例中，如果步骤 3'在步骤 4 之后发送，则步骤 4 和步骤 6 不携带 AI 模型信息。

10 还需要注意，如果步骤 3 发生并选择/切换/更新新的 AI 模型，并不意味着立即激活新模型，而是可以通过步骤 9 的 RRC 重配置指示进行激活。具体地，步骤 9 携带激活/去激活配置信息，包括至少以下之一指示模型的激活/去激活起始时间、偏置、持续时间和结束时间等。可选地，步骤 9 仅携带模型激活/去激活指示信息。当终端接收到该指示信息后，自行进行模型激活/去激活。

需要注意，上述步骤仅为示例，并不限制本发明的范围。在不脱离本发明精神和范围的情况下，可以对步骤进行各种修改和变化。

15 所描述的步骤（信令/框）的顺序不旨在被理解为限制，并且可以跳过或以任何顺序组合任意数量的所描述的步骤（信令/框）以实现方法或可替换的方法。

基站（如 gNB）在切换过程中确定快速模型控制的过程

图 10 图示了根据本公开的一些实施例的在切换过程期间的快速模型控制过程的信令图。

20 如图 10 所示，展示了潜在切换过程中的快速模型控制。对于本例，模型控制（例如模型选择、模型切换、模型更新等）位于终端侧。

示例性的，步骤 1：终端接收传统测量配置，该配置通过专用信令（例如 RRCReconfiguration 或 RRCResume 等）提供。

示例性的，步骤 2：终端根据步骤 1 中的测量配置向第一基站（如 Source gNB, S-gNB）发送测量上报消息。

25 示例性的，步骤 3：基于步骤 2 中的上报消息，第一基站发起切换(如 handover, HO)决策。

示例性的，步骤 4：第一基站检查配置的 AI/ML 模型区域信息列表，如果候选第二基站（如 Target gNB, T-gNB）在配置的 AI/ML 模型区域信息列表范围内，则无需执行模型控制。此时，第一基站将 AI/ML 模型信息发送给第二基站，并可在步骤 8 的切换请求消息中携带。如果候选第二基站不在配置的 AI/ML 模型区域信息列表范围内，则第一基站执行步骤 5，步骤 6，步骤 7。

30

示例性的，步骤 5：终端接收用于模型控制的消息，该消息用于指示终端当前的 AI/ML 模型需要进行模型控制。该消息至少包括以下内容之一：

位置信息：用于指示当前第二基站的位置信息；

5 AI/ML 模型区域信息：用于指示第二基站对应的 AI/ML 模型区域信息，例如 AI/ML 模型区域指示符、AI/ML 模型区域指示符列表、AI 模型信息、位置信息或 AI/ML 模型区域信息与模型信息的映射规则/关系；还可以用于帮助终端选择/切换 AI/ML 模型；

指示符：用于指示需要执行模型控制，和/或当前 AI/ML 模型不再适用；

事由：用于指示 AI/ML 模型更改或未更改的原因。

示例性的，步骤 6：终端根据步骤 5 执行模型控制。

10 示例性的，步骤 7：终端向第一基站发送带有模型控制结果的消息，，例如 AI/ML 模型信息；或指示符，用于指示 AI/ML 模型已经更改；

示例性的，步骤 8：第一基站向第二基站发送切换请求消息，该消息可能携带 AI/ML 模型信息，AI/ML 模型信息表示适用于第二基站的 AI 模型。此步骤可在终端和第二基站之间提前进行模型识别。

15 示例性的，步骤 9：第二基站执行接纳模型控制。

示例性的，步骤 10：第二基站向第一基站发送关于切换请求的确认消息，其中携带第二基站的信息和终端的资源信息。

示例性的，步骤 11：第一基站向终端发送带有切换命令的消息（如 RRCReconfiguration 消息等），其中携带用于访问第二基站的信息，例如小区 ID。

20 需要注意，在某些示例中，步骤 4-7 可以与步骤 3、步骤 8、步骤 9-11 同时执行。

还需要注意，在某些示例中，步骤 4 之后，第一基站检查配置的 AI/ML 模型区域信息列表。如果候选第二基站在配置的 AI/ML 模型区域信息列表范围内，表明终端侧 AI/ML 适用于第二基站，则无需执行模型控制。此时，第一基站也可以向终端发送指示消息，该消息用于指示终端，终端侧 AI/ML 适用于第二基站，可以不必执行模型控制，或第一基站也可以向终端
25 发送传统的消息，该消息包含指示符，该指示符用于指示终端，终端侧 AI/ML 适用于第二基站，可以不必执行模型控制

需要注意，上述步骤仅为示例，并不限制本发明的范围。在不脱离本发明精神和范围的情况下，可以对步骤进行各种修改和变化。

所描述的步骤（信令/框）的顺序不旨在被理解为限制，并且可以跳过或以任何顺序组合任意数量的所描述的步骤（信令/框）以实现方法或可替换的方法。
30

图 11 图示了根据本公开的一些实施例的在切换过程期间的快速模型控制过程的信令图。

如图 11 所示，展示了潜在切换过程中的快速模型控制。对于本例，模型控制（例如模型选择、模型切换、模型更新等）位于基站（如 gNB）端。

示例性的，步骤 1：终端接收传统测量配置，该传统测量配置通过专用信令（例如 RRCReconfiguration、RRCResume 等）提供。

示例性的，步骤 2：终端根据步骤 1 中的测量配置向第一基站（如 Source gNB, S-gNB）发送测量报告消息。

示例性的，步骤 3：基于步骤 2 中的报告消息，第一基站发起切换 HO 决策。

示例性的，步骤 4：第一基站检查配置的 AI/ML 模型区域信息列表，如果候选第二基站（如 Target gNB, T-gNB）在配置的 AI/ML 模型区域信息列表范围内，则无需执行模型控制。第一基站执行步骤 6，此时，第一基站可以将 AI/ML 模型信息发送给第二基站，并可以在步骤 6 的切换请求消息中携带。如果候选第二基站不在配置的 AI/ML 模型区域信息列表范围内，则第一基站执行步骤 4'，步骤 5。

示例性的，步骤 4'：或者，第一基站发起 AI/ML 模型区域信息更新过程，具体细节在实施例 2 中显示。

示例性的，步骤 5：终端根据步骤 4' 执行模型控制，例如选择/切换一个或多个适合的 AI/ML 模型。

示例性的，步骤 6：第一基站向第二基站发送切换请求消息，该消息可能携带模型信息，模型信息表示适用于第二基站的 AI/ML 模型。此步骤可以在终端和第二基站之间提前进行模型识别。

示例性的，步骤 7：第二基站执行接纳模型控制。

示例性的，步骤 8：第二基站向第一基站发送关于切换请求的确认消息，其中携带第二基站的信息和终端的资源信息。

示例性的，步骤 9：第一基站向终端发送带有切换命令的消息（如 RRCReconfiguration 消息等），其中携带用于访问第二基站的信息，例如小区 ID。

需要注意，如果在步骤 5 中选择/切换/更新了新的 AI 模型，并不意味着立即激活新模型。可以通过步骤 7 的切换请求消息进行激活指示。具体地，步骤 7 携带激活/去激活配置信息，包括指示模型的激活/去激活起始时间、偏置、持续时间和结束时间等。可选地，仅携带模型激活/去激活指示信息。当收到该指示信息后，自行进行模型激活/去激活。

需要注意，模型信息可以通过步骤 9 的消息（如 RRCReconfiguration 消息等）传输，也可

可以在步骤 4 之后通过专用的 RRC 消息、MAC CE 或 DCI 发送。

需要注意，上述步骤仅为示例，并不限制本发明的范围。在不脱离本发明精神和范围的情况下，可以对步骤进行各种修改和变化。

所描述的步骤（信令/框）的顺序不旨在被理解为限制，并且可以跳过或以任何顺序组合任意数量的所描述的步骤（信令/框）以实现方法或可替换的方法。

本公开在上述实施方式中描述网络架构中终端与网元组件之间的通信的示例，其主要出于示例目的而不是限制性的。

所描述的步骤（信令/框）的顺序不旨在被解释为限制，并且能够跳过或以任何顺序组合任何数量的所描述的步骤（信令/框）以实现方法或替代方法。通常，能够使用软件、固件、硬件（例如，固定逻辑电路）、手动处理或其任意组合来实现本文中描述的组件、模块、方法和操作的任何一个。可以在存储在计算机处理系统本地和/或远程的计算机可读存储器上的可执行指令的一般场境中描述示例方法的一些操作，并且实施方式能够包括软件应用、程序、函数等。替代地或另外，本文中描述的任何功能能够至少部分地由一个或多个硬件逻辑组件执行，诸如但不限于现场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、片上系统(SoC)、复杂可编程逻辑器件(CPLD)等。

此外，本公开实施例中描述的信令传递可以以本领域已知的任何方式来实现。例如，信令传递可以为显式和/或隐式的。此外，所图示的步骤（信令/框）仅出于示例目的，而不旨在限制本申请。

可以从以下枚举的示例实施例中理解本发明的各个方面：

根据示例实施例，提供了一种无线通信系统中的人工智能/机器学习 AI/ML 模型的控制方法，包括：

所述无线通信节点获取 AI/ML 模型区域信息，无线通信节点基于所述 AI/ML 模型区域信息确定 AI/ML 模型的模型控制，所述无线通信节点为基站或终端。

根据示例实施例，所述 AI/ML 模型的模型控制包括以下各项中的至少一者：

AI/ML 模型激活、AI/ML 模型去激活、AI/ML 模型选择、AI/ML 模型切换、AI/ML 模型回退、AI/ML 模型更新、AI/ML 模型训练、AI/ML 模型重训练、或 AI/ML 模型微调。

根据示例实施例，所述 AI/ML 模型区域信息包括以下各项中的至少一者：

AI/ML 模型区域标识 ID，

位置信息，

场景信息，

和/或 AI/ML 模型信息。

根据示例实施例，所述位置信息包括以下各项中的至少一者：

小区信息、AI/ML 模型区域标识 ID、终端位置信息、基站位置信息，场景信息、控制单元 CU ID、数据单元 DU ID。

5 根据示例实施例，所述场景信息包括以下各项中的至少一者：

传输环境信息，速度信息，角度扩展信息，时延扩展信息，信噪比，信道接收功率，和/或干扰信息。

根据本公开的一个或多个实施例，所述无线通信节点能力信息包括以下各项中的至少一者：

10 计算能力，存储能力，AI/ML 模型支持能力，负载情况，和/或传输能力，AI/ML 模型支持 AI/ML 使用场景和/或 AI/ML 模型支持 AI/ML 特征。详细而言，在一些示例中例如一些特定计算能力，一些特定存储能力，AI/ML 模型支持一些特定能力，一些特定负载情况，一些特定传输能力，AI/ML 模型支持一些特定使用场景和/或 支持支持一些特定特征。

15 无线网络能力节点能力信息；指示一个无线网络节点的能力，如节点支持的 AI/ML 模型的模型操作能力、节点的硬件能力信息（如存储能力，算力，电池能力等）、节点支持的 AI/ML 模型的功能性/特征，节点支持的 AI/ML 模型的数据，节点支持的 AI/ML 模型的场景等。

根据本公开的一个或多个实施例，所述 AI/ML 模型信息包括以下各项中的至少一者：

AI/ML 模型 ID 信息，和/或 AI/ML 模型元信息。

根据示例实施例，所述 AI/ML 模型信息包括以下各项中的至少一者：

AI/ML 模型 ID 信息，和/或 AI/ML 模型元信息。

20 根据示例实施例，所述 AI/ML 模型区域信息通过以下各项中的至少一者传输：

非接入层 NAS 消息、无线电资源控制 RRC 消息、介质接入控制控制元素 MAC-CE 消息、下行链路控制信息 DCI 消息、上行链路信道、下行链路信道、系统消息、N2 消息和/或 Xn 消息。

25 根据示例实施例，由所述无线通信节点基于所述 AI/ML 模型区域信息确定 AI/ML 模型的模型控制包括以下各项中的至少一者：

如果所述无线通信节点的位置信息属于所述 AI/ML 模型区域信息中所指示的位置信息，则所述无线通信节点不执行 AI/ML 模型的模型控制；

如果所述无线通信节点的位置信息不属于所述 AI/ML 模型区域信息中所指示的位置信息，则所述无线通信节点执行 AI/ML 模型的模型控制；

30 如果所述无线通信节点的 AI/ML 模型信息属于所述 AI/ML 模型区域信息中所指示的

AI/ML 模型信息，则所述无线通信节点不执行 AI/ML 模型的模型控制；

如果所述无线通信节点的 AI/ML 模型信息不属于所述 AI/ML 模型区域信息中所指示的 AI/ML 模型信息，则所述无线通信节点执行 AI/ML 模型的模型控制；

根据示例实施例，所述执行 AI/ML 模型的模型控制包括以下各项中的至少一者：

- 5 AI/ML 模型激活、AI/ML 模型去激活、AI/ML 模型选择、AI/ML 模型切换、AI/ML 模型回退、AI/ML 模型更新、AI/ML 模型训练、AI/ML 模型重训练、或 AI/ML 模型微调、AI/ML 模型的模型控制的结果上报，和/或 AI/ML 模型的模型控制的响应。

根据示例实施例，当所述无线通信节点为终端时，所述终端获取所述 AI/ML 模型区域配置信息，所述终端根据所述 AI/ML 模型区域信息执行所述 AI/ML 模型的模型控制。

- 10 根据示例实施例，所述终端获取所述 AI/ML 模型区域配置信息包括以下各项中的至少一者：

AI/ML 模型信息、事由、指示符、终端 ID、和/或位置信息。

根据示例实施例，当所述无线通信节点为基站时，所述基站获取所述 AI/ML 模型区域配置信息，所述基站根据所述 AI/ML 模型区域信息执行所述 AI/ML 模型的模型控制。

- 15 根据示例实施例，所述基站获取所述 AI/ML 模型区域配置信息包括以下各项中的至少一者：

AI/ML 模型信息、事由、指示符、基站 ID，和/或位置信息。

根据示例实施例，第一无线通信节点向第二无线通信节点发送的所述 AI/ML 模型区域信息请求消息和/或响应消息包括以下各项中的至少一者：

- 20 AI/ML 模型区域标识 ID，
位置信息，
场景信息，
无线通信节点能力信息，和/或
AI/ML 模型信息。

- 25 根据示例实施例，所述无线通信节点发送所述 AI/ML 模型区域信息的请求消息，包括周期性发送或事件触发。

根据示例实施例，在所述周期性发送的情况下，所述 AI/ML 模型区域信息响应于周期性配置而进行更新，其中，所述周期性配置包括所述更新的持续时间、起始时间、和/或结束时间。

- 30 根据示例实施例，在所述方式为事件触发的情况下，所述 AI/ML 模型区域信息响应于事件触发配置而进行更新。

根据示例实施例，所述周期性发送和/或所述事件触发相应的周期性配置和/或事件触发配置是预配置的、和/或配置的、和/或固定的。

根据示例实施例，所述 AI/ML 模型区域信息请求消息和/或响应消息通过以下至少一项承载：

- 5 非接入层 NAS 消息、无线电资源控制 RRC 消息、介质接入控制控制元素 MAC-CE 消息、下行链路控制信息 DCI 消息、上行链路信道、下行链路信道、系统消息、N2 消息和/或 Xn 消息。

根据示例实施例，提供了一种用于终端的无线通信方法，包括：

所述终端从第一基站获取 AI/ML 模型区域信息，所述 AI/ML 模型区域信息用于所述终端

- 10 确定 AI/ML 模型的模型控制，

其中，所述终端基于所述 AI/ML 模型区域信息确定是否执行所述 AI/ML 模型的模型控制。

根据示例实施例，还包括：

所述终端从所述第一基站获取 AI/ML 模型区域配置消息。

- 15 根据示例实施例，如果第二基站不在所述 AI/ML 模型区域的情况下，所述终端确定执行所述 AI/ML 模型的模型控制。

根据示例实施例，所述执行 AI/ML 模型的模型控制包括在所述终端与所述第一基站之间更新所述 AI/ML 模型区域信息。

根据示例实施例，还包括：

- 20 所述终端将针对所述 AI/ML 模型区域配置消息的上报消息以及所述执行所述 AI/ML 模型的模型控制的结果发送至所述第一基站，所述第一基站将获取的执行所述 AI/ML 模型的模型控制的结果发送至所述第二基站。

根据示例实施例，提供了一种用于第一基站的无线通信方法，包括：

所述第一基站从所述终端获取 AI/ML 模型的模型控制上报消息，所述 AI/ML 模型的模型控制上报消息包括所述 AI/ML 模型的模型控制的结果；

- 25 所述第一基站将所述第二基站的小区切换请求以及所述 AI/ML 模型的模型控制的结果发送至所述第二基站。

根据示例实施例，提供了一种用于第一基站的无线通信方法，包括：

所述第一基站根据终端反馈的 AI/ML 模型的模型控制消息，确定第二基站是否在 AI/ML 模型区域信息中；

- 30 如果所述第二基站不在所述 AI/ML 模型区域信息中，所述第一基站执行 AI/ML 模型的模

型控制。

根据示例实施例，所述 AI/ML 模型的模型控制包括在所述第一基站与所述第二基站之间更新所述 AI/ML 模型区域信息。

5 根据示例实施例，还包括：所述第一基站将所述第二基站的小区切换请求以及所述模型控制的结果发送至所述第二基站。

根据示例实施例，还包括：如果所述第二基站在所述 AI/ML 模型区域信息中，所述第一基站将所述 AI/ML 模型区域信息中所包含的 AI/ML 模型信息以及至所述第二基站的小区切换请求发送至所述第二基站。

10 根据示例实施例，提供了一种芯片，包括：处理器，其被配置为调用和运行存储在存储器中的计算机程序，以使得其中安装有所述芯片的设备执行示例实施例中任一个的方法。

图 12 是根据本公开的实施例的用于无线通信的示例系统 700 的框图。可以使用任何适当配置的硬件和/或软件将本文描述的实施例实现到系统中。图 12 图示系统 700，包括射频 (RF) 电路 710、基带电路 720、处理单元 730、存储器/储存器 740、显示器 750、相机 760、传感器 770 和输入/输出 (I/O) 接口 780，如图所示彼此耦合。

15 处理单元 730 可以包括电路，例如但不限于一个或多个单核或多核处理器。处理器可以包括通用处理器和专用处理器的任何组合，例如图形处理器和应用处理器。处理器可以与存储器/储存器耦合，并且被配置为执行存储在存储器/储存器中的指令，以能够实现在系统上运行的各种应用和/或操作系统。RF 电路 710、基带电路 720、处理单元 730、存储器/储存器 740、显示器 750、相机 760、传感器 770 和 I/O 接口 780 是系统 700 中的公知元件，例如但不限于膝
20 上型计算设备、平板计算设备、上网本、超极致笔电、智能电话等。此外，作为软件产品的指令可以存储在计算机中的可读存储介质中。计算机中的软件产品存储在存储介质中，包括用于计算设备（诸如个人计算机、服务器或网络设备）的多个命令，以运行本公开的实施例所公开的所有或一些步骤。存储介质包括 USB 盘、移动硬盘、只读存储器 (ROM)、随机存取存储器 (RAM)、软盘或能够存储程序代码的其它类型的介质。

25 本公开的实施例是可以在 3GPP 规范中采用的技术/过程的组合以创建最终产品。

虽然已经结合被认为是最实用和优选的实施例描述了本公开，但是应当理解，本公开不限于所公开的实施例，而是旨在覆盖在不脱离所附权利要求的最广泛解释的范围的情况下所做出的各种布置。

权 利 要 求 书

1. 一种无线通信系统中的人工智能/机器学习 AI/ML 模型的控制方法，包括：

5 所述无线通信节点获取 AI/ML 模型区域信息，无线通信节点基于所述 AI/ML 模型区域信息确定 AI/ML 模型的模型控制，所述无线通信节点为基站或终端。

2. 根据权利要求 1 所述的方法，其中，所述 AI/ML 模型的模型控制包括以下各项中的至少一者：

AI/ML 模型激活、AI/ML 模型去激活、AI/ML 模型选择、AI/ML 模型切换、AI/ML 模型回退、AI/ML 模型更新、AI/ML 模型训练、AI/ML 模型重训练、或 AI/ML 模型微调。

10 3. 根据权利要求 1 所述的方法，其中，所述 AI/ML 模型区域信息包括以下各项中的至少一者：

AI/ML 模型区域标识 ID，

位置信息，

场景信息，

15 无线通信节点能力信息，

和/或 AI/ML 模型信息。

4. 根据权利要求 3 所述的方法，其中，所述位置信息包括以下各项中的至少一者：

小区信息、AI/ML 模型区域标识 ID、终端位置信息、基站位置信息，场景信息、控制单元 CU ID、数据单元 DU ID。

20 5. 根据权利要求 3 所述的方法，其中，所述场景信息包括以下各项中的至少一者：

传输环境信息，速度信息，角度扩展信息，时延扩展信息，信噪比，信道接收功率，和/或干扰信息。

6. 根据权利要求 3 所述的方法，其中，所述无线通信节点能力信息包括以下各项中的至少一者：

25 计算能力，存储能力，AI/ML 模型支持能力，负载情况，传输能力，AI/ML 模型支持 AI/ML 使用场景和/或 AI/ML 模型支持 AI/ML 特征。

7. 根据权利要求 3 所述的方法，其中，所述 AI/ML 模型信息包括以下各项中的至少一者：

AI/ML 模型 ID 信息，和/或 AI/ML 模型元信息。

8. 根据权利要求 1 所述的方法，其中，所述 AI/ML 模型区域信息通过以下各项中的至少一者传输：

非接入层 NAS 消息、无线电资源控制 RRC 消息、介质接入控制控制元素 MAC-CE 消息、下行链路控制信息 DCI 消息、上行链路信道、下行链路信道、系统消息、N2 消息和/或 Xn 消息。

9. 根据权利要求 3 所述的方法，其中，由所述无线通信节点基于所述 AI/ML 模型区域信息确定 AI/ML 模型的模型控制包括以下各项中的至少一者：

如果所述无线通信节点的位置信息属于所述 AI/ML 模型区域信息中所指示的位置信息，则所述无线通信节点不执行 AI/ML 模型的模型控制；

如果所述无线通信节点的位置信息不属于所述 AI/ML 模型区域信息中所指示的位置信息，则所述无线通信节点执行 AI/ML 模型的模型控制；

10. 如果所述无线通信节点的 AI/ML 模型信息属于所述 AI/ML 模型区域信息中所指示的 AI/ML 模型信息，则所述无线通信节点不执行 AI/ML 模型的模型控制；

如果所述无线通信节点的 AI/ML 模型信息不属于所述 AI/ML 模型区域信息中所指示的 AI/ML 模型信息，则所述无线通信节点执行 AI/ML 模型的模型控制。

11. 根据权利要求 9 所述的方法，其中，所述执行 AI/ML 模型的模型控制包括以下各项中的至少一者：

AI/ML 模型激活、AI/ML 模型去激活、AI/ML 模型选择、AI/ML 模型切换、AI/ML 模型回退、AI/ML 模型更新、AI/ML 模型训练、AI/ML 模型重训练、或 AI/ML 模型微调、AI/ML 模型的模型控制的结果上报，和/或 AI/ML 模型的模型控制的响应。

12. 根据权利要求 10 所述的方法，其中，当所述无线通信节点为终端时，所述终端获取所述 AI/ML 模型区域配置信息，所述终端根据所述 AI/ML 模型区域信息执行所述 AI/ML 模型的模型控制。

13. 根据权利要求 11 所述的方法，其中，所述终端获取所述 AI/ML 模型区域配置信息包括以下各项中的至少一者：

AI/ML 模型信息、事由、指示符、终端 ID、和/或位置信息。

14. 根据权利要求 10 所述的方法，其中，当所述无线通信节点为基站时，所述基站获取所述 AI/ML 模型区域配置信息，所述基站根据所述 AI/ML 模型区域信息执行所述 AI/ML 模型的模型控制。

15. 根据权利要求 13 所述的方法，其中，所述基站获取所述 AI/ML 模型区域配置信息包括以下各项中的至少一者：

AI/ML 模型信息、事由、指示符、基站 ID，和/或位置信息。

15. 根据权利要求 1 所述的方法，其中，第一无线通信节点向第二无线通信节点发送的所述 AI/ML 模型区域信息请求消息和/或响应消息包括以下各项中的至少一者：

AI/ML 模型区域标识 ID，

位置信息，

5 场景信息，

无线通信节点能力信息，和/或

AI/ML 模型信息。

16. 根据权利要求 15 所述的方法，其中，所述无线通信节点发送所述 AI/ML 模型区域信息的请求消息，包括周期性发送或事件触发。

10 17. 根据权利要求 16 所述的方法，其中，在所述周期性发送的情况下，所述 AI/ML 模型区域信息响应于周期性配置而进行更新，其中，所述周期性配置包括所述更新的持续时间、起始时间、和/或结束时间。

18. 根据权利要求 16 所述的方法，其中，在所述方式为事件触发的情况下，所述 AI/ML 模型区域信息响应于事件触发配置而进行更新。

15 19. 根据权利要求 16 所述的方法，其中，所述周期性发送和/或所述事件触发相应的周期性配置和/或事件触发配置是预配置的、和/或配置的、和/或固定的。

20. 根据权利要求 15 所述的方法，所述 AI/ML 模型区域信息请求消息和/或响应消息通过以下至少一项承载：

非接入层 NAS 消息、无线电资源控制 RRC 消息、介质接入控制控制元素 MAC-CE 消息、

20 下行链路控制信息 DCI 消息、上行链路信道、下行链路信道、系统消息、N2 消息和/或 Xn 消息。

21. 一种用于终端的无线通信方法，包括：

所述终端从第一基站获取 AI/ML 模型区域信息，所述 AI/ML 模型区域信息用于所述终端确定 AI/ML 模型的模型控制，

25 其中，所述终端基于所述 AI/ML 模型区域信息确定是否执行所述 AI/ML 模型的模型控制。

22. 根据权利要求 21 所述的方法，还包括：

所述终端从所述第一基站获取 AI/ML 模型区域配置消息。

23. 根据权利要求 22 所述的方法，其中，如果第二基站不在所述 AI/ML 模型区域的情况下，所述终端确定执行所述 AI/ML 模型的模型控制。

30 24. 根据权利要求 23 所述的方法，其中，所述执行 AI/ML 模型的模型控制包括在所述终

端与所述第一基站之间更新所述 AI/ML 模型区域信息。

25. 根据权利要求 21 所述的方法，还包括：

所述终端将针对所述 AI/ML 模型区域配置消息的上报消息以及所述执行所述 AI/ML 模型的模型控制的结果发送至所述第一基站，所述第一基站将获取的执行所述 AI/ML 模型的模型控制的结果发送至所述第二基站。

26. 一种终端，所述终端被配置成执行权利要求 21 至 25 中任一项所述的方法。

27. 一种用于第一基站的无线通信方法，包括：

所述第一基站从所述终端获取 AI/ML 模型的模型控制上报消息，所述 AI/ML 模型的模型控制上报消息包括所述 AI/ML 模型的模型控制的结果；

10 所述第一基站将所述第二基站的小区切换请求以及所述 AI/ML 模型的模型控制的结果发送至所述第二基站。

28. 一种用于第一基站的无线通信方法，包括：

所述第一基站根据终端反馈的 AI/ML 模型的模型控制消息，确定第二基站是否在 AI/ML 模型区域信息中；

15 如果所述第二基站不在所述 AI/ML 模型区域信息中，所述第一基站执行 AI/ML 模型的模型控制。

29. 根据权利要求 28 所述的方法，其中，所述 AI/ML 模型的模型控制包括在所述第一基站与所述第二基站之间更新所述 AI/ML 模型区域信息。

30. 根据权利要求 29 所述的方法，还包括：

20 所述第一基站将所述第二基站的小区切换请求以及所述模型控制的结果发送至所述第二基站。

31. 根据权利要求 28 所述的方法，还包括：

如果所述第二基站在所述 AI/ML 模型区域信息中，所述第一基站将所述 AI/ML 模型区域信息中所包含的 AI/ML 模型信息以及至所述第二基站的小区切换请求发送至所述第二基站。

25 32. 一种芯片，包括：

处理器，其被配置为调用和运行存储在存储器中的计算机程序，以使得其中安装有所述芯片的设备执行权利要求 1 至 31 中任一项所述的方法。

附 图

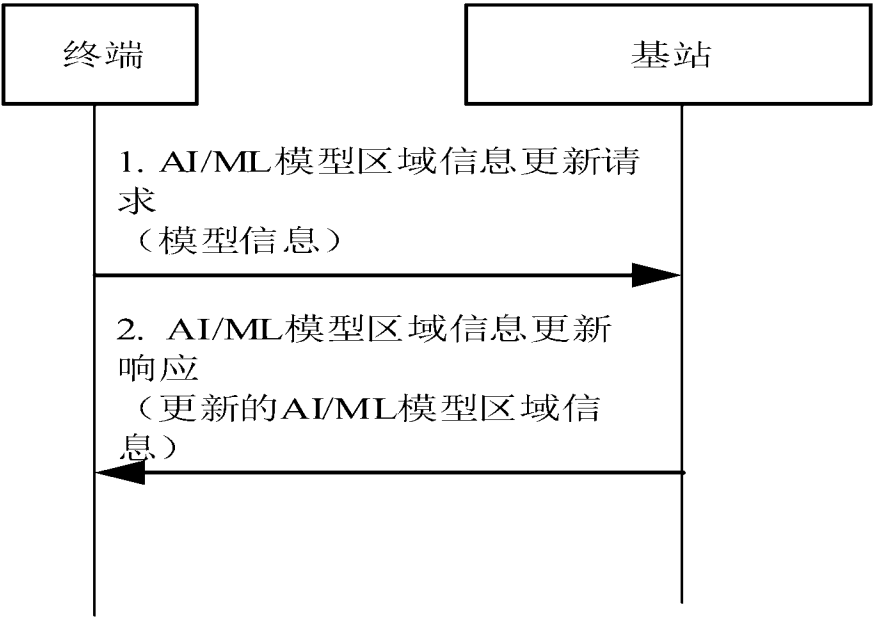


图 1

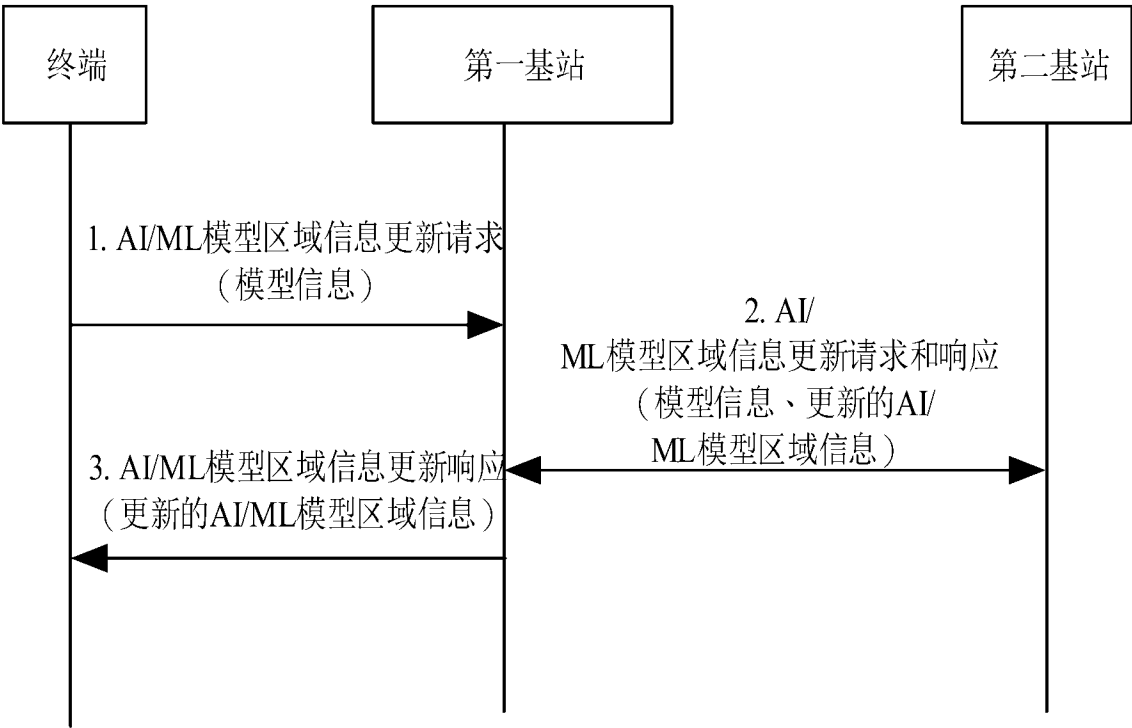


图 2

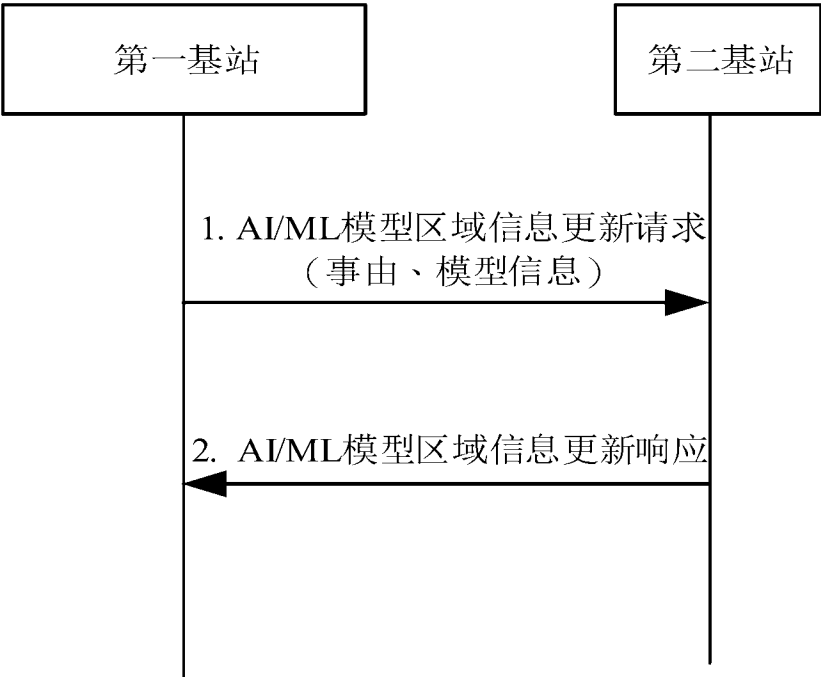


图 3

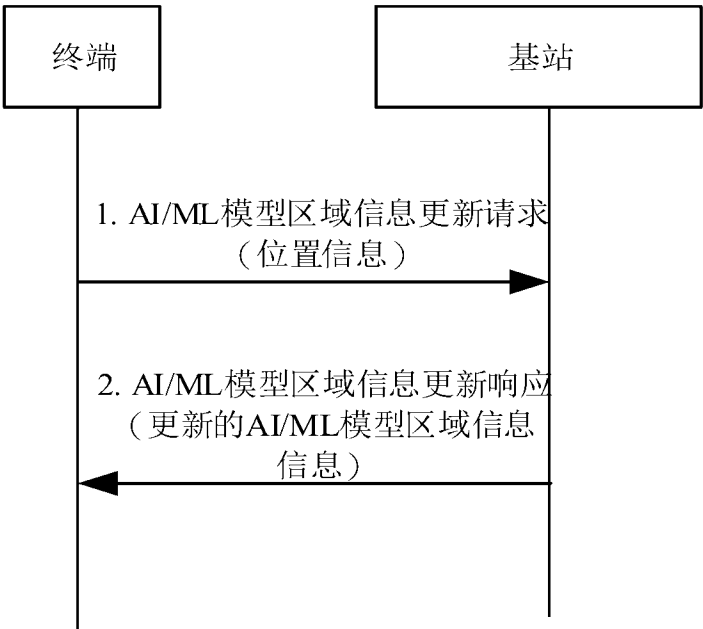


图 4

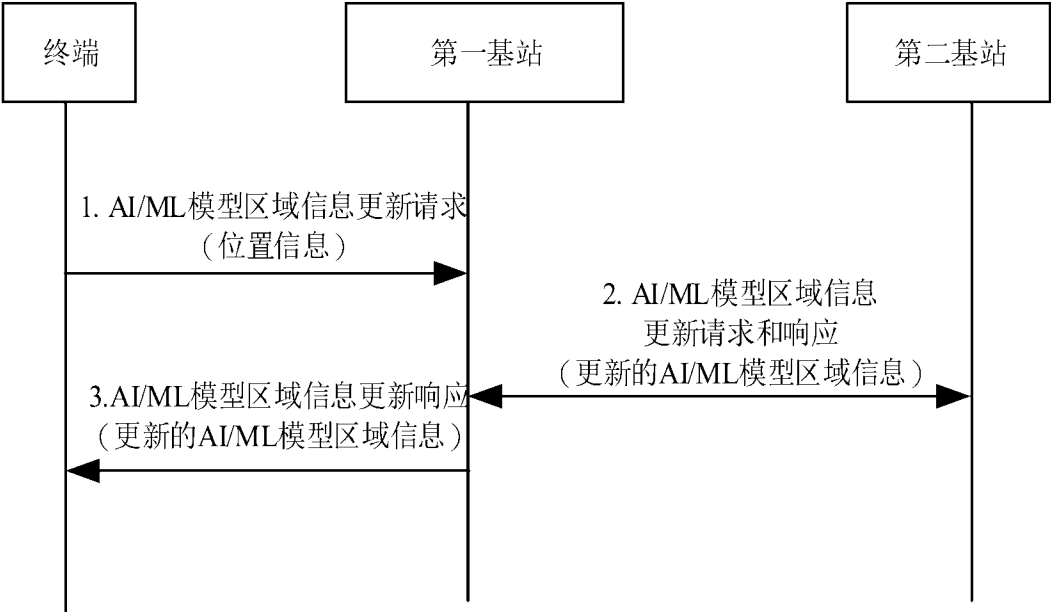


图 5

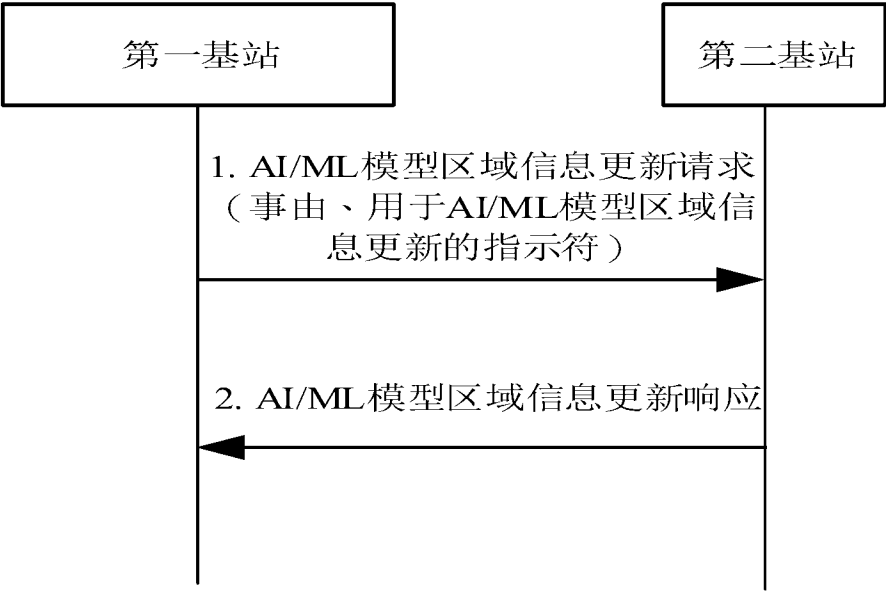


图 6

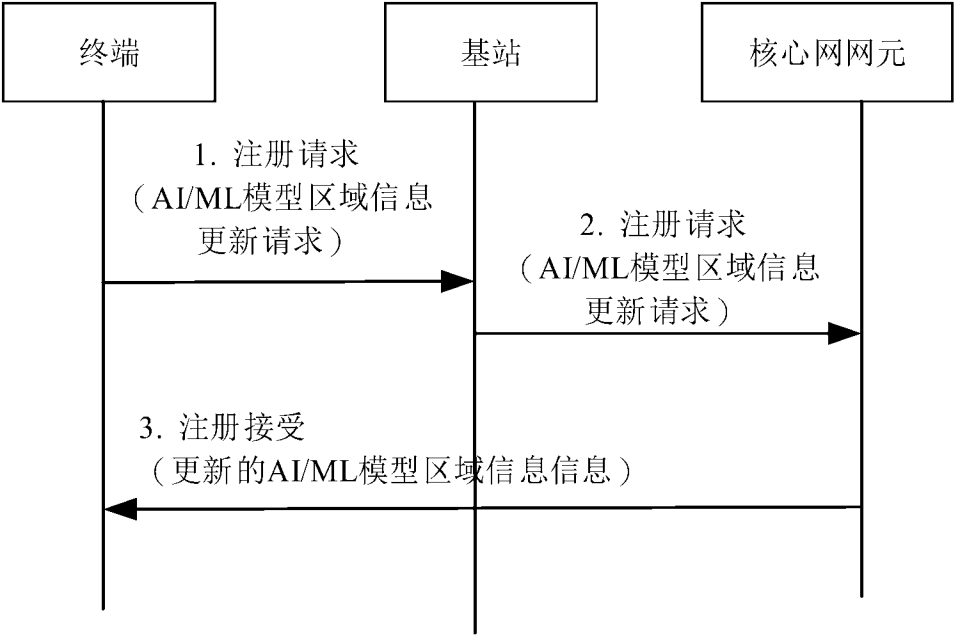


图 7

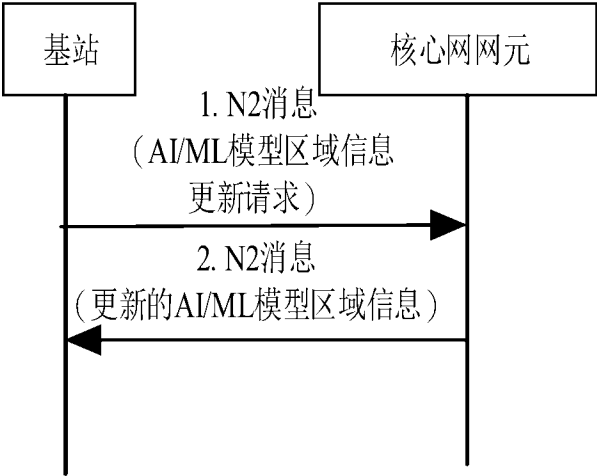


图 8

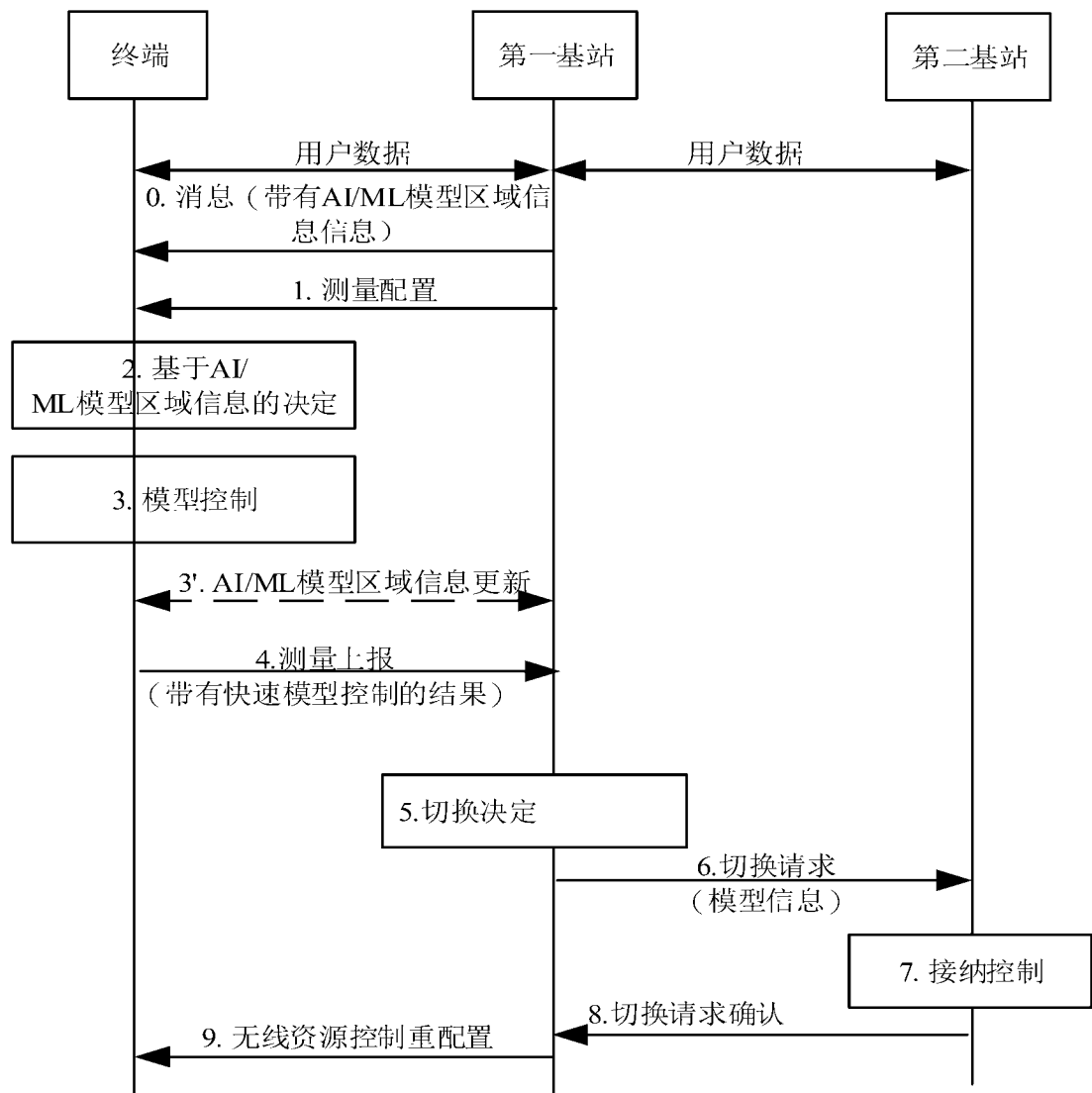


图 9

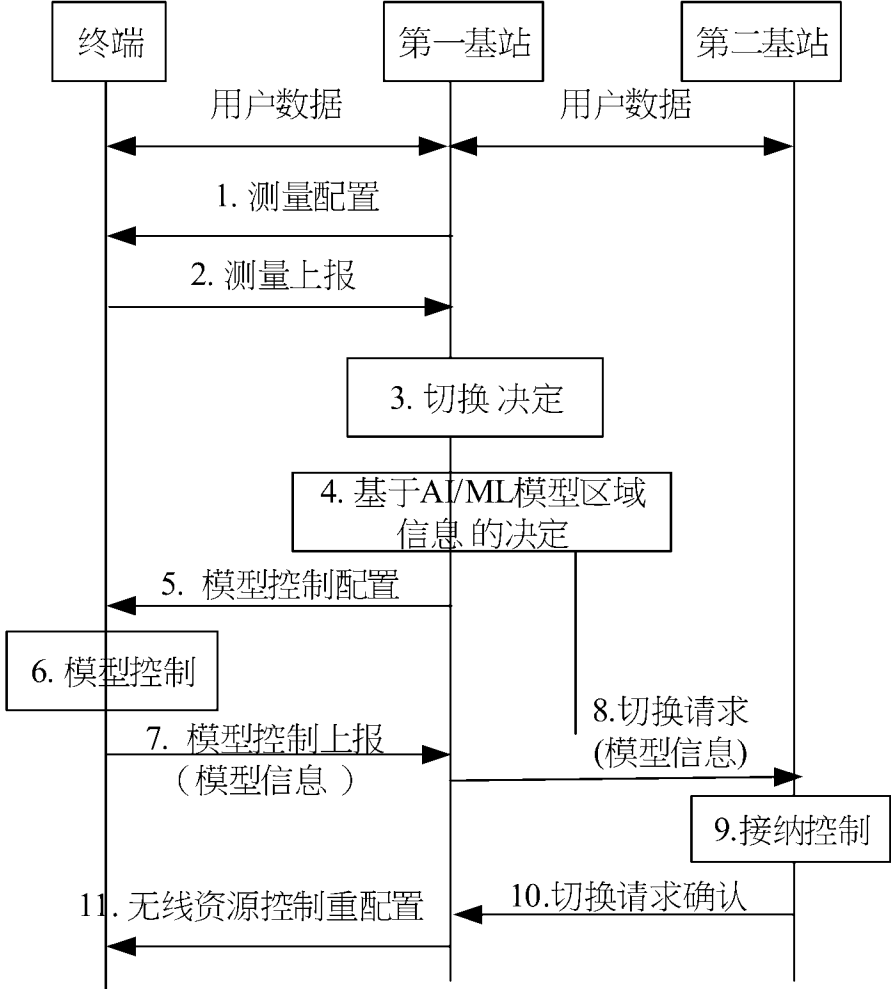


图 10

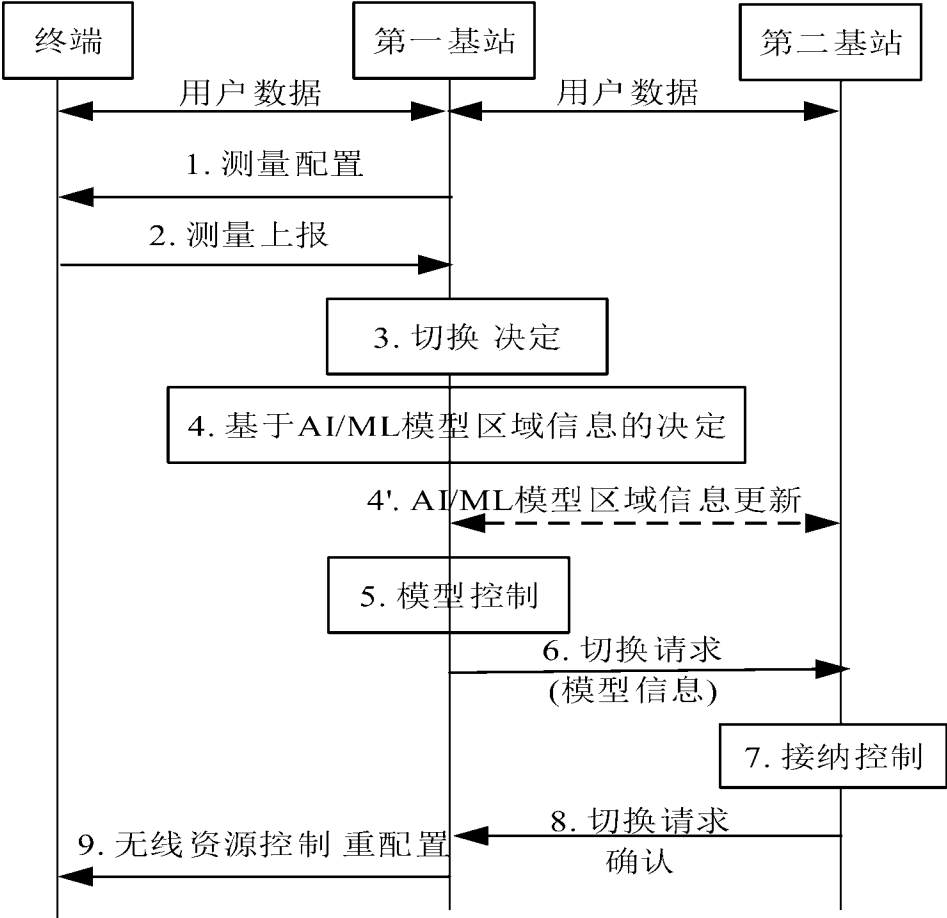


图 11

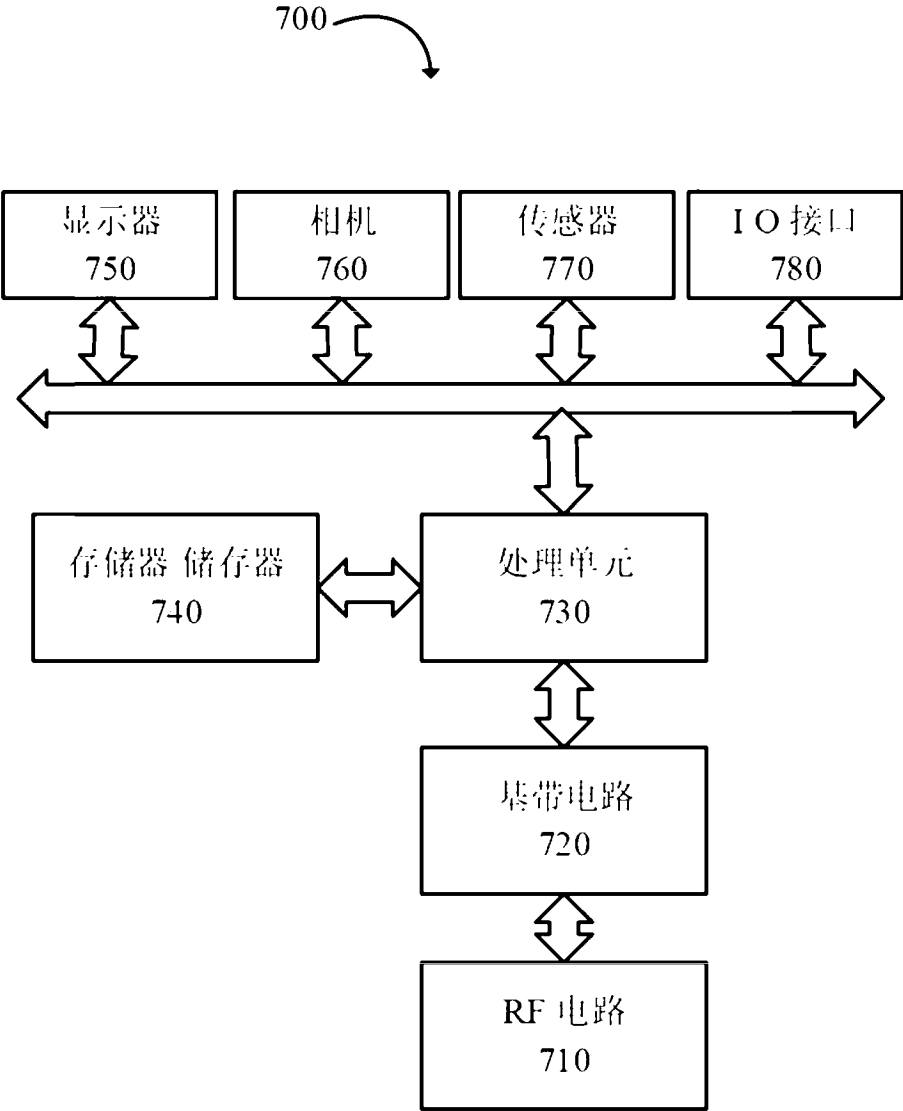


图 12

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/123047

A. CLASSIFICATION OF SUBJECT MATTER H04L41/16(2022.01)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: H04L, H04W, H04Q 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) CNTXT, WPABS, WPABSC, ENTXTC, VEN, VCN, 3GPP: 人工智能, 机器学习, 模型, 区域, 控制, 切换, 更新, AI, ML, model, handover, update, ID		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2023133897 A1 (SHENZHEN TCL NEW TECHNOLOGY CO., LTD.) 20 July 2023 (2023-07-20) description, paragraphs [0004]-[0070]	1-32
X	MEDIATEK INC. "Discussion on RAN2 Aspects for LCM" 3GPP TSG-RAN WG2 Meeting #120, R2-2211610, 18 November 2022 (2022-11-18), sections 1-3	1-32
A	CN 109479309 A (SAMSUNG ELECTRONICS CO., LTD.) 15 March 2019 (2019-03-15) entire document	1-32
A	US 2023093963 A1 (QUALCOMM INC.) 30 March 2023 (2023-03-30) entire document	1-32
☐ 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 29 May 2024		Date of mailing of the international search report 19 June 2024
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/CN2023/123047

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)		Publication date (day/month/year)
WO	2023133897	A1	20 July 2023	None		
CN	109479309	A	15 March 2019	EP	3908045 A1	10 November 2021
				EP	3469846 A1	17 April 2019
				EP	3469846 A4	26 May 2021
				US	2018020382 A1	18 January 2018
				US	10349318 B2	09 July 2019
				KR	20210107582 A	01 September 2021
				KR	102334508 B1	06 December 2021
				KR	20180007697 A	23 January 2018
				KR	102295185 B1	02 September 2021
				WO	2018012904 A1	18 January 2018
				US	2022240132 A1	28 July 2022
				US	11736988 B2	22 August 2023
				US	2019320354 A1	17 October 2019
				US	11330480 B2	10 May 2022
				US	2019320353 A1	17 October 2019
				US	11223980 B2	11 January 2022
US	2023093963	A1	30 March 2023	None		

A. 主题的分类

H04L41/16(2022.01)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: H04L, H04W, H04Q

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))

CNTEXT, WPABS, WPABSC, ENTXTC, VEN, VCN, 3GPP: 人工智能, 机器学习, 模型, 区域, 控制, 切换, 更新, AI, ML, model, handover, update, ID

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	WO 2023133897 A1 (SHENZHEN TCL NEW TECHNOLOGY CO., LTD.) 2023年7月20日 (2023 - 07 - 20) 说明书第[0004]-[0070]段	1-32
X	MEDIATEK INC. "Discussion on RAN2 aspects for LCM" 3GPP TSG-RAN WG2 Meeting #120, R2-2211610, 2022年11月18日 (2022 - 11 - 18), 第1-3节	1-32
A	CN 109479309 A (三星电子株式会社) 2019年3月15日 (2019 - 03 - 15) 全文	1-32
A	US 2023093963 A1 (QUALCOMM INCORPORATED) 2023年3月30日 (2023 - 03 - 30) 全文	1-32

☐ 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

2024年5月29日

国际检索报告邮寄日期

2024年6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

杨盈霄

电话号码 (+86) 010-53961572

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/CN2023/123047

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)
WO	2023133897	A1	2023年7月20日	无	
CN	109479309	A	2019年3月15日	EP	3908045 A1 2021年11月10日
				EP	3469846 A1 2019年4月17日
				EP	3469846 A4 2021年5月26日
				US	2018020382 A1 2018年1月18日
				US	10349318 B2 2019年7月9日
				KR	20210107582 A 2021年9月1日
				KR	102334508 B1 2021年12月6日
				KR	20180007697 A 2018年1月23日
				KR	102295185 B1 2021年9月2日
				WO	2018012904 A1 2018年1月18日
				US	2022240132 A1 2022年7月28日
				US	11736988 B2 2023年8月22日
				US	2019320354 A1 2019年10月17日
				US	11330480 B2 2022年5月10日
				US	2019320353 A1 2019年10月17日
				US	11223980 B2 2022年1月11日
US	2023093963	A1	2023年3月30日	无	