(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 26 日 (26.06.2025)

WIPO | PCT

(10) 国际公布号
WO 2025/130470 A1

(51) 国际专利分类号:
G06F 9/48 (2006.01)

(21) 国际申请号: PCT/CN2024/132767

(22) 国际申请日: 2024 年 11 月 18 日 (18.11.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202311769643.8 2023年12月20日 (20.12.2023) CN

(71) 申请人: 华 为 技 术 有 限 公 司 **(HUAWEI TECHNOLOGIES CO., LTD.)** [CN/CN]; 中 国 广 东 省 深 圳 市 龙 岗 区 坂 田 华 为 总 部 办 公 楼 518129 (CN)。

(72) 发明人:刘银萍**(LIU, Yinping)**; 中国广东省深圳市龙岗区坂田华为总部办公楼 518129 (CN)。 刘宗惠 **(LIU, Zonghui)**; 中国广东省深圳市龙岗区坂田华为总部办公楼 518129 (CN)。

(74) 代理人: 北 京 中 博 世 达 专 利 商 标 代 理 有 限 公 司 **(BEIJING ZBSD PATENT & TRADEMARK AGENT LTD.)**; 中 国 北 京 市 海 淀 区 交 大 东 路 31号11号楼8层 100044 (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:** METHOD AND APPARATUS FOR DETERMINING TASK EXECUTION POLICY

(54) 发明名称: 确定任务执行策略的方法和装置

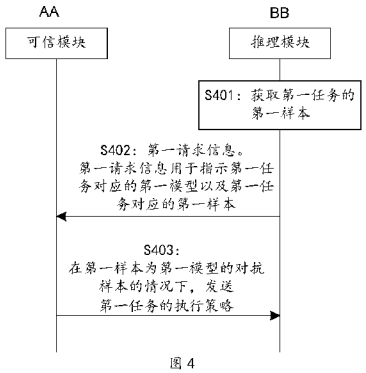


图 4

S401 Acquire a first example of a first task
S402 First request information, which is used for indicating a first model corresponding to the first task and the first example corresponding to the first task
S403 When the first example is an adversarial example of the first model, send an execution policy for the first task
AA Trusted module
BB Inference module

(57) **Abstract:** The present application relates to the technical field of artificial intelligence. Provided are a method and apparatus for determining a task execution policy. In the method, a trusted module can receive first request information for indicating a first model corresponding to a first task and a first example corresponding to the first task, and when the first example is an adversarial example of the first model, the trusted module can send an execution policy for the first task. The execution policy for the first task is used for indicating a mode of converting the first example into a non-adversarial example. Since the execution policy for the first task can indicate the mode of converting the first example into the non-adversarial example, the problem of a model inference error caused by the adversarial example can be ameliorated, so that the error rate in model inference is reduced.

WO 2025/130470 A1

[见续页]

(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: 本申请提供确定任务执行策略的方法和装置，涉及人工智能技术领域。在该方法中，可信模块可以接收用于指示第一任务对应的第一模型以及第一任务对应的第一样本的第一请求信息，在第一样本为第一模型的对抗样本的情况下，可信模块可以发送第一任务的执行策略。其中，第一任务的执行策略用于指示将第一样本变更为非对抗样本的方式。由于第一任务的执行策略可以指示将第一样本变更为非对抗样本的方式，所以可以改善因对抗样本导致的模型推理出错的问题，从而降低模型推理出错率。

确定任务执行策略的方法和装置

本申请要求于 2023 年 12 月 20 日提交国家知识产权局、申请号为 202311769643.8、发明名称为“确定任务执行策略的方法和装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及人工智能（artificial intelligence，AI）技术领域，尤其涉及确定任务执行策略的方法和装置。

背景技术

随着人工智能（artificial intelligence，AI）技术地发展，越来越多的业务可以使用模型，如 AI 模型或机器学习（machine learning，ML）模型进行推理。模型对输入数据有高度敏感性，所以即使输入数据（即样本）被施加难以察觉的细微扰动，也可能会导致模型推理出错误的结果。因此，如何降低模型推理的出错率是目前亟需解决的问题。

发明内容

本申请提供确定任务执行策略的方法和装置，可以降低模型推理的出错率。

为达到上述目的，本申请的采用如下技术方案：

第一方面，提供了一种确定任务执行策略的方法，该方法可以由可信模块执行。这里的可信模块既可以指可信模块本身，也可以指可信模块中实现该方法的处理器、模块、逻辑节点、芯片、或芯片系统等。

该方法包括：接收第一请求信息；第一请求信息用于指示第一任务对应的第一模型以及第一任务对应的第一样本；在第一样本为第一模型的对抗样本的情况下，发送第一任务的执行策略，第一任务的执行策略用于指示将第一样本变更为非对抗样本的方式。

基于上述第一方面提供的方法，由于第一任务的执行策略可以指示将第一样本变更为非对抗样本的方式，所以该方法可以改善因对抗样本导致的模型推理出错的问题，从而降低模型推理出错率。

在一种可能的实现方式中，将第一样本变更为非对抗样本的方式包括：更换第一模型；或者，更换第一样本；或者，对第一样本执行第一操作。

基于上述可能的实现方式，将第一样本变更为非对抗样本的方式包括更换第一模型时，可以通过更换第一模型的方式将第一样本变更为非对抗样本；将第一样本变更为非对抗样本的方式包括更换第一样本时，可以通过更换第一样本的方式将第一样本变更为非对抗样本；将第一样本变更为非对抗样本的方式包括对第一样本执行第一操作时，可以通过对第一样本执行第一操作将第一样本变更为非对抗样本。

在一种可能的实现方式中，上述方法还包括：获取第一指示信息，第一指示信息用于指示第一任务是否有备用样本，和/或，第一指示信息用于指示第一任务是否有备用模型；根据第一指示信息确定第一任务的执行策略。

基于上述可能的实现方式，可以根据第一指示信息确定第一任务的执行策略，即根据是否有备用样本，和/或，第一任务是否有备用模型确定第一任务的执行策略。

在一种可能的实现方式中，根据第一指示信息确定第一任务的执行策略，包括：第一指示信息指示第一任务有备用样本，第一任务的执行策略指示更换第一样本；或者，第一指示信息指示第一任务有备用模型，第一任务的执行策略指示更换第一模型；第一指示信息指示第一任务没有备用样本也没有备用模型，第一任务的执行策略指示对第一样本执行第一操作。

基于上述可能的实现方式，在第一指示信息指示第一任务有备用样本，第一任务的执行策略指示更换第一样本的情况下，可以使第一模型根据备用样本得到输出结果，避免第一模型根据第一样本输出错误结果；在第一指示信息指示第一任务有备用模型，第一任务的执行策略指示更换第一模型的情况下，可以使第一任务的备用模型根据第一样本得到输出结果，避免第一模型根据第一样本输出错误结果；在第一指示信息指示第一任务没有备用样本也没有备用模型，第一任务的执行策略指示对第一样本执行第一操作的情况下，可以使第一模型根据执行第一操作之后得到的样本得到输出结果，避免第一模型根据第一样本输出错误的结果。

在一种可能的实现方式中，第一操作包括以下至少一项：特征压缩、样本去噪或数据平滑处理。

基于上述可能的实现方式，可以通过对第一样本进行特征压缩、样本去噪或数据平滑等方法，剔除导致第一模型推理出错的样本数据，使执行第一操作之后得到的样本变更为第一模型的非对抗样本。

在一种可能的实现方式中，第一任务的执行策略指示更换第一样本，该方法还包括：获取第一任务对应的第二样本；根据第一模型确定第二样本是否是对抗样本。

基于上述可能的实现方式，还可以获取第一任务对应的第二样本，并根据第一模型确定第二样本是否是对抗样本，以便降低第一模型推理的出错率。

在一种可能的实现方式中，第一任务的执行策略指示更换第一模型，该方法还包括：获取第一任务对应的第二模型；根据第二模型确定第一样本是否是对抗样本。

基于上述可能的实现方式，还可以获取第一任务对应的第二模型，并根据第二模型确定第一样本是否是对抗样本，以便降低第一任务的推理结果的出错率。

在一种可能的实现方式中，上述方法还包括：接收第一策略信息，第一策略信息指示第一任务的执行策略。

基于上述可能的实现方式，可信模块可以根据第一策略信息确定第一样本变更为非对抗样本的方式。

在一种可能的实现方式中，上述方法还包括：获取第一任务的鲁棒性需求，第一任务的鲁棒性需求指示需要检测第一任务对应的第一样本是否是对抗样本；根据第一模型确定第一样本是否为对抗样本。

基于上述可能的实现方式，可信模块可以根据第一任务的鲁棒性需求，确定是否需要检测第一样本是否是第一模型的对抗的样本。当第一任务的鲁棒性需求指示检测第一任务对应的第一样本是否是对抗样本时，可信模块可以确定需要检测第一样本是否是第一模型的对抗的样本。当第一任务的鲁棒性需求指示不检测第一任务对应的第一样本是否是对抗样本时，可信模块可以确定不需要检测第一样本是否是第一模型的对抗的样本。

在一种可能的实现方式中，第一任务的鲁棒性需求还指示第一任务对应的第一模型的输出结果的鲁棒性要求；根据第一模型确定第一样本是否为对抗样本，包括：根据第一模型和鲁棒性要求确定第一样本是否为对抗样本。

基于上述可能的实现方式，可以结合鲁棒性要求，如具体不稳定变化的阈值更灵活地根据第一模型确定第一样本是否为对抗样本。

第二方面，提供了一种确定任务执行策略的方法，该方法可以由推理模块执行。这里的推理模块既可以指推理模块本身，也可以指推理模块中实现该方法的处理器、模块、逻辑节点、芯片、或芯片系统等。

该方法包括：获取第一任务的第一样本；发送第一请求信息；第一请求信息指示第一任务的第一样本和第一任务的第一模型；接收第一任务的执行策略，第一任务的执行策略指示用于将第一样本变更为非对抗样本的方式。

基于上述第二方面提供的方法，由于第一任务的执行策略可以指示将第一样本变更为非对抗样本的方式，所以该方法可以改善因对抗样本导致的模型推理出错的问题，从而降低模型推理出错率。

在一种可能的实现方式中，将第一样本变更为非对抗样本的方式包括：更换第一模型；或者，更换第一样本；或者，对第一样本执行第一操作。

基于上述可能的实现方式，将第一样本变更为非对抗样本的方式包括更换第一模型时，推理模块可以通过更换第一模型的方式将第一样本变更为非对抗样本；将第一样本变更为非对抗样本的方式包括更换第一样本时，推理模块可以通过更换第一样本的方式将第一样本变更为非对抗样本；将第一样本变更为非对抗样本的方式包括对第一样本执行第一操作时，推理模块可以根据对第一样本执行第一操作后的样本进行推理。

在一种可能的实现方式中，上述方法还包括：发送第一指示信息，第一指示信息用于指示第一任务是否有备用样本，和/或，第一指示信息用于指示第一任务是否有备用模型，第一指示信息用于确定第一任务的执行策略。

基于上述可能的实现方式，可以使得接收到第一指示信息的装置，如可信模块，根据第一指示信息指示确定第一任务的执行策略。

在一种可能的实现方式中，第一指示信息指示第一任务有备用样本，第一任务的执行策略指示更换第一样本；或者，第一指示信息指示第一任务有备用模型，第一任务的执行策略指示更换第一模型；第一指示信息指示第一任务没有备用样本也没有备用模型，第一任务的执行策略指示对第一样本执行第一

操作。

基于上述可能的实现方式，在第一指示信息指示第一任务有备用样本，第一任务的执行策略指示更换第一样本的情况下，可以使第一模型根据备用样本得到输出结果，避免第一模型根据第一样本输出错误结果；在第一指示信息指示第一任务有备用模型，第一任务的执行策略指示更换第一模型的情况下，可以使第一任务的备用模型根据第一样本得到输出结果，避免第一模型根据第一样本输出错误结果；在第一指示信息指示第一任务没有备用样本也没有备用模型的情况下，第一任务的执行策略指示对第一样本执行第一操作，可以使第一模型根据执行第一操作后的第一样本得到输出结果，避免第一模型根据第一样本输出错误的结果。

在一种可能的实现方式中，第一操作包括以下至少一项：特征压缩、样本去噪或数据平滑处理。

基于上述可能的实现方式，通过对第一样本进行特征压缩、样本去噪或数据平滑等方法，可以剔除导致第一模型推理出错的样本数据，使执行第一操作之后得到的样本变更为第一模型的非对抗样本。

在一种可能的实现方式中，第一任务的执行策略指示更换第一模型，上述方法还包括：接收来自通信节点的第二指示信息，第二指示信息用于指示第一任务的备用模型。

基于上述可能的实现方式，第一任务的执行策略指示更换第一模型时，可以根据第二指示信息，获取第一任务的备用模型。

在一种可能的实现方式中，第一任务的执行策略指示更换第一模型时，上述方法还包括：向通信节点发送第一样本；接收来自通信节点的推理结果，该推理结果用于指示根据第一样本和第一任务的备用模型进行推理得到的结果。

基于上述可能的实现方式，可以通过向通信节点发送第一样本，以获取根据第一样本和备用模型进行推理得到的结果。

在一种可能的实现方式中，上述方法还包括：接收第一任务的鲁棒性需求，第一任务的鲁棒性需求指示需要检测第一任务对应的样本是否是对抗样本。

基于上述可能的实现方式，可以使得推理模块根据第一任务的鲁棒性需求确定是否发送第一请求信息。

在一种可能的实现方式中，第一任务的鲁棒性需求还指示第一任务的模型的输出结果的鲁棒性要求。

基于上述可能的实现方式，可以确定第一任务的模型的输出结果的鲁棒性要求。

在一种可能的实现方式中，第一任务的执行策略还用于指示第一样本为对抗样本。

基于上述可能的实现方式，可以根据第一任务的执行策略确定第一样本变更为非对抗样本的方式。

第三方面，提供了一种通信装置用于实现上述方法。该通信装置可以为上述第一方面中的可信模块；或者，该通信装置可以为上述第二方面中的推理模块。该通信装置包括实现上述方法相应的模块、单元、或手段（means），该模块、单元、或means可以通过硬件实现，软件实现，或者通过硬件执行相应的软件实现。该硬件或软件包括一个或多个与上述功能相对应的模块或单元。

结合上述第三方面，在一种可能的实现方式中，该通信装置可以包括处理模块和接口模块。该处理模块，可以用于实现上述任一方面及其任意可能的实现方式中的处理功能。该处理模块例如可以为处理器。该接口模块，也可以称为接口单元，用以实现上述任一方面及其任意可能的实现方式中的发送和/或接收功能。该接口模块可以由接口电路，收发机，收发器或者通信接口构成。

结合上述第三方面，在一种可能的实现方式中，接口模块包括发送模块和接收模块，分别用于实现上述任一方面及其任意可能的实现方式中的发送和接收功能。

第四方面，提供了一种通信装置，包括：处理器；该处理器用于与存储器耦合，并读取存储器中的指令之后，根据该指令执行如上述任一方面所述的方法。该通信装置可以为上述第一方面中的可信模块；或者，该通信装置可以为上述第二方面中的推理模块。

结合上述第四方面，在一种可能的实现方式中，该通信装置还包括存储器，该存储器，用于保存程序指令和数据。可选的，该存储器与上述处理器集成在一起；或者，该存储器独立于该处理器。

结合上述第四方面，在一种可能的实现方式中，该通信装置为芯片或芯片系统。可选的，该通信装置是芯片系统时，可以由芯片构成，也可以包含芯片和其他分立器件。

第五方面，提供了一种通信装置，包括：处理器和接口电路；接口电路，用于接收计算机程序或指令并传输至处理器；处理器用于执行所述计算机程序或指令，以使该通信装置执行如上述任一方面所述的方法。该通信装置可以为上述第一方面中的可信模块；或者，该通信装置可以为上述第二方面中的推

理模块。

结合上述第五方面，在一种可能的实现方式中，该通信装置为芯片或芯片系统。可选的，该通信装置是芯片系统时，可以由芯片构成，也可以包含芯片和其他分立器件。

第六方面，提供了一种计算机可读存储介质，该计算机可读存储介质中存储有指令，当其在计算机上运行时，使得计算机可以执行上述任一方面所述的方法。

第七方面，提供了一种包含指令的计算机程序产品，当其在计算机上运行时，使得计算机可以执行上述任一方面所述的方法。

第八方面，提供了一种通信系统，该通信系统包括用于执行上述第一方面所述的方法的可信模块、以及用于执行上述第二方面所述的方法的推理模块。

其中，第三方面至第八方面中任一种可能的实现方式所带来的技术效果可参见上述第一方面至第二方面中任一方面或任一方面中不同可能的实现方式所带来的技术效果，此处不再赘述。

可以理解的是，在方案不矛盾的前提下，上述各个方面中的方案均可以结合。

附图说明

图 1A 为本申请提供的一种分类模型的稳定区域和不稳定区域的示意图；

图 1B 为本申请提供的一个图片分类模型受到对抗样本攻击的示意图；

图 1C 为本申请提供的一个回归模型受到对抗样本攻击的示意图；

图 1D 为本申请提供的波束管理用例的 AI 流程示意图；

图 1E 为本申请提供的一个波束管理模型受到对抗样本攻击的示意图；

图 2A 为本申请提供的通信系统架构示意图一；

图 2B 为本申请提供的通信系统架构示意图二；

图 2C 为本申请提供的通信系统架构示意图三；

图 2D 为本申请提供的无线接入网络（radio access network，RAN）域（domain）的 AI 工作流程示意图；

图 2E 为本申请提供的操作、管理和维护（operations,administration and maintenance，OAM）域的 AI 工作流程示意图；

图 2F 为本申请提供的通信系统架构示意图四；

图 2G 为本申请提供的通信系统架构示意图五；

图 3 为本申请提供的通信装置的硬件结构示意图；

图 4 为本申请提供的通信方法流程示意图一；

图 5 为本申请提供的通信方法流程示意图二；

图 6 为本申请提供的通信装置的结构示意图。

具体实施方式

在模型的推理阶段（即根据模型输入的推理样本得到输出结果的阶段），如果模型不够鲁棒（robust），在噪声、测量误差、恶意修改等多种因素的作用下，推理样本可能成为模型的对抗样本，导致模型的推理出错。在 ML 或 AI 领域，“鲁棒”可理解为使模型能够抵御一定的恶意攻击的性能，让模型具备鲁棒性是为了在最大程度上保证模型的输出结果不受攻击的影响而改变。

上文中的对抗样本是指在模型的推理阶段特别设计一种针对模型的样本，如通过在原始样本中加入微小但精心设计的扰动，来误导模型，从而使模型得到错误的输出结果。

对抗样本存在的原因根植于模型的固有缺陷，固有缺陷是指模型天然存在的不稳定区域，下面以分类模型为例进行说明。如图 1A 展示了一种分类模型的稳定区域和不稳定区域的示意图，图 1A 中每个圆形和正方形分别对应模型根据不同输入得到的输出结果，其中圆形可以对应模型输出的分类结果为“A”的点，正方形可以对应模型输出的分类结果为“B”的点。可以理解的，距离分类边界越近的点，对输入数据变化越敏感，即距离分类边界越近的点，当输入被施加扰动时，越容易使模型输出错误的结果。分类边界附近的区域可以叫做不稳定区域（也可以叫做不鲁棒区域），远离分类边界的区域可以叫做稳定区域。当模型的分类边界附近不稳定区域范围越大时，分类边界附近推理样本的微小扰动将导致错误输出，对抗样本存在的概率也越大。针对上文中的分类边界，可以有如下理解：以分类模型为例，可以根据概率（置信度）对若干种不同的可能结果进行决策，不同的结果对应不同概率的决策门限，例如 50%以上可以

决策为A，那么50%附近的数据，对结果产生的干扰程度较严重。

因为模型对数据非常敏感，所以这些微小的扰动虽然对于用户来说通常是不可察觉的，却可以使模型输出错误的结果，例如做出错误的分类结果或错误的数值结果等。下面以示例 1 和示例 2 为例进行说明。

示例 1，如图 1B 所示，以一个图片分类模型受到对抗样本攻击的为例进行说明。原图片为熊猫，如果把原图片作为该图片分类模型的输入，该图片分类模型基于原图片输出的分类结果是“熊猫”（置信度是 57.7%）。如果在原图片基础上叠加一些精心设计的肉眼不可见的噪声（例如，噪声可以是进行 0.007 加权）后，得到对抗样本图片，该图片分类模型基于对抗样本图片输出的分类结果是“长臂猿”（置信度 99.3%）。

示例 2，如图 1C 所示，以一个回归模型受到对抗样本攻击为例进行说明。其中，回归模型可以是对统计关系进行定量描述的一种数学模型。该回归模型可以用表达式 f(x)表示，输入数据为 x 时，该回归模型输出 y。如果对 x 施加一些干扰，输入数据变为（x+Δx1），此时该回归模型输出 y’。对 x 施加一些其他干扰，输入数据也可能变为（x+Δx2），此时该回归模型输出 y’’。y’’ 和 y’ 是错误的输出数据。

可以理解的，上述模型推理过程中受到对抗样本影响的事件也叫做对抗攻击，具体的，对抗攻击可理解为对目标机器学习模型的原输入的推理样本施加轻微扰动以生成对抗样本、来欺骗目标模型的过程。

因为模型受到对抗攻击后可能产生难以想象的后果，所以 AI 的鲁棒性问题需要谨慎对待，事实上国际很多相关组织都非常重视此类问题，例如，针对高风险 AI 系统，欧盟 AI 法案草案中明确提出 AI 系统需要进行一定的鲁棒性增强，免受投毒、对抗等攻击。欧盟 AI 法案（ACT ARTICLE 15）中给出了 AI 鲁棒性的要求：应解决 AI 特有的脆弱性，包括在适当的情况下采取措施预防、检测、响应、解决并控制试图操纵训练数据集（“数据投毒（data poisoning）”）的攻击，或者用于训练的预训练组件（“模型投毒（model poisoning）”），旨在导致模型出错（“对抗样本”或“模型闪避（model evasion）”），机密泄露（confidentiality attacks）或其他模型缺陷，这些都可能导致有害的决策。

下面以波束管理用例为例对 AI 流程进行说明。如图 1D 所示，以波束管理用例的模型为例，对具体步骤说明如下，其中 S101~S106 对应模型的训练阶段，S107~S111 对应模型的推理阶段。

S101：NR 中的 RAN 节点向终端发送全波束扫描的指示。相应的，终端接收全波束扫描的指示。

其中，全波束是指全方向的波束，并且，全方向具体指示的角度值可以由 RAN 节点设置，例如，全方向可以是 180°或 360°等角度。全波束扫描可以是 32 波束或 64 波束或 256 波束的波束扫描，具体波束数量是由 RAN 节点配置的，本示例以 64 波束为全波束为例进行说明。

S102：终端获取波束扫描的训练模型的数据。

对于波束扫描用例，终端为获取波束扫描的训练模型的推理样本，可以对全波束进行测量，分别得到每个波束的参考信号接收功率（reference signal receiving power，RSRP）。所有波束的标识（Identity，ID）和对应的 RSRP 可以作为训练模型的数据。

S103：终端向 RAN 节点发送训练模型的数据。相应的，RAN 节点接收来自终端的训练模型的数据。

S104：RAN 节点生成或训练模型。

可以理解的，RAN 节点利用来自终端的训练模型的数据对进行模型生成或训练，输出数据是 64 波束中 5 个最优波束的发生概率。

S105：RAN 节点向终端发送稀疏波束扫描的指示。相应的，终端接收来自 RAN 节点的稀疏波束扫描的指示。

本例中，以 16 波束作为稀疏波束的波束数量进行说明。

S106：终端向 RAN 节点发送稀疏波束对应的 RSRP。相应的，RAN 节点接收来自终端的稀疏波束对应的 RSRP。

应理解 S105~S106 这两个步骤的组合可以多次执行，用于是 RAN 节点获取多个稀疏波束对应的样本，这些样本可以作为训练样本对模型进行训练。

S107：RAN 节点利用模型推理出 5 个最优波束。

可以理解的，RAN 节点利用 S104 生成的模型推理出 64 波束中的 5 个最优波束，得到 5 个最优波束的波束 ID。

另外，最优波束的数量可以通过预设置等方式设定，本例以 5 个最优波束为例进行说明，最优波束的数量也可以是其他数量，不作限制。RAN 节点从 64 波束中推理出的 5 个最优波束的一种示意图，可以参

见S107旁的网格图，通过网格图中5个黑色的方块可以指示出5个最优波束。

S108：RAN节点向终端发送基于5个最优波束进行二阶段扫描的指示。相应的，终端接收来自RAN节点的基于5个最优波束进行二阶段扫描的指示。

S109：终端测量波束RSRP。

终端测量RSRP，从5个最优波束中选择出最优波束。

S110：终端向RAN节点发送最优波束的指示。相应的，RAN节点接收来自终端的最优波束的指示。

可以理解的，最优波束的指示可以包括最优波束的ID。

S111：RAN节点基于最优波束发送信号。相应的，终端接收来自RAN节点的基于最优波束发送的信号。

可以理解的，在上述波束管理用例的模型推理阶段，结合图1E，RAN节点可以向多个终端分别获取稀疏波束扫描结果，从而得到多个样本（参见图1E中的图左一），通过多个样本训练出一个模型，模型的推理过程可以用表达式f(x)表示，模型的输出结果是64波束中的5个最佳波束。在模型的推理过程中，如果终端测量波束的RSRP时未出现误差，即模型输入的推理样本的数据是正确的，则模型输出的是正确结果（可以参考图1E中的图右一）。如果终端测量波束的RSRP时存在误差，且误差大到使样本落在模型的不稳定区域时，模型的输出很可能发生错误，推理样本就成为模型的对抗样本，输出错误结果（参见图1E中的图左二）。

由此可见，在模型的推理阶段，模型会因为输入数据的细微扰动出现推理出错的情况。

为了解决上述问题，本申请提供一种确定任务执行策略的方法。在该方法中，可信模块可以接收第一请求信息，其中第一请求信息用于指示第一任务对应的第一模型以及第一任务对应的第一样本，根据第一模型确定第一样本是否为对抗样本。在第一样本为对抗样本的情况下，可信模块可以确定第一任务的执行策略，第一任务的执行策略用于指示将第一样本变更为非对抗样本的方式。通过上述方案，可信模块可以确定第一样本是否是对抗样本，如果第一样本为对抗样本，可信模块可以确定将第一样本变更为非对抗样本的方式，从而避免因对抗样本导致模型推理出错，模型推理出错率会得到改善。

下面结合附图对本申请的实施方式进行详细描述。

可以理解的，本申请提供的方法可用于各种AI系统。下面以图2A所示AI系统20为例，对本申请提供的方法进行描述。图2A仅为示意图，并不构成对本申请提供的技术方案的适用场景的限定。

如图2A所示，为本申请提供的AI系统20的架构示意图。图2A中，AI系统20可以包括可信模块201以及与可信模块201连接的推理模块202。可选的，AI系统20还包括管理模块203。

其中，可信模块201可以用于获取第一任务对应的第一模型以及第一任务对应的第一样本，根据第一模型确定第一样本是否为对抗样本，在第一样本为对抗样本的情况下，确定第一任务的执行策略。第一任务的执行策略用于指示将第一样本变更为非对抗样本的至少一种方式。可信模块201还用于将第一任务的执行策略发送给推理模块202，以便推理模块202根据该执行策略执行第一任务。

管理模块203可以用于向可信模块201发送第一任务的鲁棒性需求，以便可信模块201根据该第一任务的鲁棒性需求确定是否需要检测第一任务对应的样本是否是对抗样本。

在一种可能的实现方式中，图2A所示的AI系统20可以应用于3GPP网络中，或者开放无线接入网络（open radio access network，ORAN）架构等，本申请实施例对此不作具体限定。下面进行具体阐述。

示例性的，图2A所示的AI系统20可以应用于图2B所示的通信网络架构。图2B所示的通信网络可以按照功能划分为RAN域，以及RAN域/跨域管理服务（management service，MnS）消费者（consumer）等。其中，RAN域包括域管理功能（例如，域管理功能可以是OAM）和RAN节点。RAN节点具备AI/ML推理功能和可信AI/ML管理功能。AI/ML推理功能可以具备推理模块202的功能，可信AI/ML管理功能具备可信模块201的功能，域管理功能可以具备管理模块203的功能。其中，RAN域也叫做接入网域。

在图2B中，RAN域通过北向接口向RAN域/跨域MnS消费者发送数据，RAN域/跨域管理服务消费者通过南向接口向RAN域发送数据。

示例性的，图2A所示的AI系统20还可以应用于图2C所示的通信网络架构。图2C区别于图2B之处在于，图2C中，域管理功能具备AI/ML推理功能、可信AI/ML管理功能。其中，AI/ML推理功能可以具备推理模块202的功能，可信AI/ML管理功能具备可信模块201的功能，域管理功能可以具备管理模块203的功能。可以理解的，除了图2B或图2C所述的两种通信系统架构之外，还可以有其他的通信系统架构，例如，RAN节点可以具备AI/ML推理功能，域管理功能可以具备可信AI/ML管理功能。或者，RAN节点可以具备可信AI/ML管理功能，域管理功能可以具备AI/ML推理功能。本申请对AI/ML推理功能和可信

AI/ML 管理功能在 3GPP 网络中部署的位置不予限制。

通过上述实现方式，可以灵活地部署 AI/ML 推理功能和可信 AI/ML 管理功能，AI/ML 推理功能和可信 AI/ML 管理功能都可以为不同的接口提供多样的服务，满足服务化接口定义的需求。下面对 RAN 节点和域管理功能相关的 AI 工作流程分别进行介绍。

以图 2B 为例，RAN 节点的 AI 工作流程主要分为数据收集（data collection）、训练（training）以及推理（inference）、使用（actor）等阶段，如图 2D 所示。首先，RAN 节点可以进行数据收集，收集得到的数据可以作为训练数据用于模型训练，也可以作为训练完成后的模型的输入数据，进行模型推理。模型的输出数据可以被应用于相应的场景，在模型的使用过程中得到的其他不同于输入数据的数据还可以反馈到数据收集模块。上述过程可以循环进行，不再赘述。

以图 2C 为例，OAM（即图 2C 中的域管理功能）的 AI 工作流程可以分为训练、部署、推理这三个阶段，其中如图 2E 所示。在训练阶段，OAM 可以对模型进行训练，并对训练完成后的模型进行测试。在模型测试的过程中，如果出现问题 OAM 可以重新进行模型训练。模型完成测试后，OAM 可以将模型部署到 RAN 节点中。模型部署完成后，OAM 可以利用这些模型进行推理。如果在推理过程中出现异常，OAM 还可以对该模型进行训练，用于进一步修正模型。可以理解的，上述修正后的模型（或其他情况的模型）完成模型测试后，也可以不经过部署（或者其他已经完成部署的模型）用于模型推理。

除了上述方式之外，图 2A 所示的 AI 系统 20 还可以应用于 ORAN 网络中。例如，AI 系统 20 还可以应用于图 2F 或图 2G 所示的通信系统架构。

可以理解的，ORAN 网络是一种开放式 RAN 架构，具有开放的标准化接口，可以独立构建各个模块，以便根据不同标准开发的蜂窝网络设备之间可以进行互操作，这样，无线网络设备提供商便可专注于提供特定的组件，而不是构建整个 RAN，从而使移动通信网络实现软件化、虚拟化、灵活化、智能化和节能。

如图 2F 或图 2G 所示，ORAN 网络的通信系统架构可以包含：非实时无线智能控制（non-real-time radio intelligent controller，Non-RT RIC）、近实时无线智能控制（near-real-time radio intelligent controller，Near-RT RIC）、ORAN 架构中央单元（ORAN-central unit，O-CU）和 ORAN 架构分布式单元（ORAN-distributed unit，O-DU）。

其中，O-CU 和 O-DU 都属于 ORAN 的节点。其中 O-CU 是 ORAN 系统下的集中单元（centralized unit，CU），主要负责非实时 L2 和无线资源控制（radio resource control，RRC）等功能。其中 O-DU 是 ORAN 系统下的分布单元（distributed unit，DU），主要负责实时 L2 功能、基带信号处理等功能。

近实时 RIC 是在无线资源管理(radio resource management，RRM)的基础之上进行了增强，综合了 RRM、切片管理、服务水平协议、AI/ML、移动边缘云计算等技术，从而实现对 RAN（即 ORAN 中的接入网部分）提供近实时的智能控制。近实时 RIC 通过 ORAN 标准化的接口分别与 O-CU、O-DU 以及非实时 RIC 连接。

非实时 RIC 位于 ORAN 的网络管理平台，执行策略管理、RAN 分析和基于 AI/ML 的功能管理。非实时 RIC 通过 ORAN 标准化的接口与近实时 RIC 连接。

近实时 RIC 和非实时 RIC 都属于实时无线智能控制（real-time radio intelligent controller，RT RIC）。一种可能的实现方式，近实时 RIC 实现对 ORAN 的节点（O-CU/O-DU）的近实时控制和优化，非实时 RIC 实现对 ORAN 节点（O-CU/O-DU）的非实时控制。例如，可以在非实时 RIC 平台中完成模型训练，并部署到近实时 RIC 平台进行推理，非实时 RIC 也可以下发执行策略至近实时 RIC。

对于 ORAN 网络，O-CU 或 O-DU 可以具备 AI/ML 推理功能，非实时 RIC 或近实时 RIC 可以具备可信 AI/ML 管理功能。图 2F 给出一种非实时 RIC 具备可信 AI/ML 管理功能的示意图。图 2G 给出一种近实时 RIC 具备可信 AI/ML 管理功能的示意图。在图 2F 或图 2G 给出的通信网络架构中，O-CU 或 O-DU 可以具备 AI/ML 推理功能。其中，可信 AI/ML 管理功能可以具备上述可信模块 201 的功能，AI/ML 推理功能可以具备上述推理模块 202 的功能。

本申请中，RAN 节点可以是一种具有无线收发功能的设备，可帮助终端实现无线接入。本申请中的 RAN 节点又可以称为 RAN 中的节点、RAN 节点或接入网设备等。RAN 节点包括但不限于：LTE 中的演进型基站（NodeB 或 eNB 或 e-NodeB，evolutional Node B），下一代 LTE 中的演进型基站（next generation eNB，ng-eNB），NR 中的基站（gNodeB 或 gNB），下一代 RAN 节点（next generation radio access network，NG-RAN）、发射点（transmitting point，TP）或收发点（transmission receiving point/transmission reception point，TRP），3GPP 后续演进的基站，下一代基站（next generation NodeB，gNB），6G 移动

通信系统中的下一代基站，未来移动通信系统中的基站，卫星，WiFi 系统中的接入节点，无线中继节点，无线回传节点，接入回传一体化（integrated access and backhaul，IAB）节点、移动交换中心非陆地通信网络（non-terrestrial network，NTN）通信系统中的 RAN 节点，即可以部署于高空平台或者卫星等。基站可以是：宏基站，微基站，微微基站，小站，中继站，或，气球站等。多个基站可以支持上述提及的同一种技术的网络，也可以支持上述提及的不同技术的网络。基站可以包含一个或多个共站或非共站的 TRP。RAN 节点还可以是 D2D 通信、车联网通信、无人机通信、机器通信中担任基站功能的设备。RAN 节点还可以是云无线接入网络（cloud radio access network，CRAN）场景下的无线控制器。RAN 节点还可以是集中单元（centralized unit，CU）、分布单元（distributed unit，DU）、CU-控制面（control plane，CP）、CU-用户面（user plane，UP）、无线单元（radio unit，RU）、具有基站功能的路边单元（road side unit，RSU）、有线接入网关或者核心网网元等。RAN 节点还可以是服务器，可穿戴设备，机器通信设备或车载设备等。例如，V2X 技术中的 RAN 节点可以为 RSU。以下以 RAN 节点为基站为例进行说明。所述多个 RAN 节点可以为同一类型的基站，也可以为不同类型的基站。基站可以与终端进行通信，也可以通过中继站与终端进行通信。终端可以与不同技术的多个基站进行通信，例如，终端可以与支持 LTE 网络的基站通信，也可以与支持 5G 网络的基站通信，还可以支持与 LTE 网络的基站以及 5G 网络的基站的双连接。

本申请中，CU 和 DU 可以是单独设置，或者也可以包括在同一个网元中，例如基带单元（baseband unit，BBU）中。RU 可以包括在射频设备或者射频单元中，例如包括在射频拉远单元（remote radio unit，RRU）、有源天线处理单元（active antenna unit，AAU）或远程射频头（remote radio head，RRH）中。可以理解的是，CU 可以划分为接入网中的 RAN 节点，也可以将 CU 划分为核心网中的 RAN 节点，在此不做限制。

在不同系统中，CU（或 CU-CP 和 CU-UP）、DU 或 RU 也可以有不同的名称，但是本领域的技术人员可以理解其含义。例如，在开放无线接入网（open radio access network，ORAN）系统中，CU 也可以称为 O-CU（开放式 CU），DU 也可以称为 O-DU，CU-CP 也可以称为 O-CU-CP，CU-UP 也可以称为 O-CU-UP，RU 也可以称为 O-RU。为描述方便，本申请中以 CU，CU-CP，CU-UP、DU 和 RU 为例进行描述。本申请中的 CU（或 CU-CP、CU-UP）、DU 和 RU 中的任一单元，可以是通过软件模块、硬件模块、或者软件模块与硬件模块结合来实现。

本申请中，对 RAN 节点的形态不作限定，用于实现 RAN 节点的功能的装置可以是 RAN 节点；也可以是能够支持 RAN 节点实现该功能的装置，例如芯片系统。该装置可以被安装在 RAN 节点中或者和 RAN 节点匹配使用。

可选的，本申请图 2A 中的各模块（例如可信模块 201、推理模块 202 或管理模块 203 等）也可以称之为通信装置，其可以是一个通用设备或者是一个专用设备，本申请对此不作具体限定。

可选的，本申请图 2A 中的各模块（例如可信模块 201、推理模块 202 或管理模块 203 等）的相关功能可以由一个设备实现，也可以由多个设备共同实现，还可以是由一个设备内的一个或多个功能模块实现，本申请对此不作具体限定。可以理解的是，上述功能既可以是硬件设备中的网络元件，也可以是在专用硬件上运行的软件功能，或者硬件与软件的结合，或者平台(例如，云平台)上实例化的虚拟化功能。

在具体实现时，本申请图 2A 中的各模块（例如可信模块 201、推理模块 202 或管理模块 203 等）都可以采用图 3 所示的组成结构，或者包括图 3 所示的部件。图 3 所示为可适用于本申请的通信装置的硬件结构示意图。该通信装置 30 包括至少一个处理器 301 和至少一个通信接口 304，用于实现本申请提供的方法。该通信装置 30 还可以包括通信线路 302 和存储器 303。

处理器 301 可以是一个通用中央处理器（central processing unit，CPU），微处理器，特定应用集成电路（application-specific integrated circuit，ASIC），或一个或多个用于控制本申请方案程序执行的集成电路。

通信线路 302 可包括一通路，在上述组件之间传送信息，例如总线。

通信接口 304，用于与其他设备或通信网络通信。通信接口 304 可以是任何收发器一类的装置，如可以是以太网接口、无线接入网（radio access network，RAN）接口、无线局域网（wireless local area networks，WLAN）接口、收发器、管脚、总线、接口电路或收发电路等。

存储器 303 可以是只读存储器（read-only memory，ROM）或可存储静态信息和指令的其他类型的静态存储设备，随机存取存储器（random access memory，RAM）或者可存储信息和指令的其他类型的动态

存储设备，也可以是电可擦可编程只读存储器（electrically erasable programmable read-only memory，EEPROM）、只读光盘（compact disc read-only memory，CD-ROM）或其他光盘存储、光碟存储（包括压缩光碟、激光碟、光碟、数字通用光碟、蓝光光碟等）、磁盘存储介质或者其他磁存储设备、或者能够用于携带或存储具有指令或数据结构形式的期望的程序代码并能够由计算机存取的任何其他介质，但不限于此。存储器可以是独立存在，通过通信线路 302 与处理器 301 相耦合。存储器 303 也可以和处理器 301 集成在一起。本申请提供的存储器通常可以具有非易失性。

其中，存储器 303 用于存储执行本申请提供的方案所涉及的计算机执行指令，并由处理器 301 来控制执行。处理器 301 用于执行存储器 303 中存储的计算机执行指令，从而实现本申请提供的方法。或者，可选的，本申请中，也可以是处理器 301 执行本申请下述提供的方法中的处理相关的功能，通信接口 304 负责与其他设备或通信网络通信，本申请对此不作具体限定。

可选的，本申请中的计算机执行指令也可以称之为应用程序代码，本申请对此不作具体限定。

本申请中的耦合是装置、单元或模块之间的间接耦合或通信连接，可以是电性，机械或其它的形式，用于装置、单元或模块之间的信息交互。

作为一种实施例，处理器 301 可以包括一个或多个 CPU，例如图 3 中的 CPU0 和 CPU1。

作为一种实施例，通信装置 30 可以包括多个处理器，例如图 3 中的处理器 301 和处理器 307。这些处理器中的每一个可以是一个单核（single-CPU）处理器，也可以是一个多核（multi-CPU）处理器。这里的处理器可以指一个或多个设备、电路、和/或用于处理数据（例如计算机程序指令）的处理核。

作为一种实施例，通信装置 30 还可以包括输出设备 305 和/或输入设备 306。输出设备 305 和处理器 301 耦合，可以以多种方式来显示信息。例如，输出设备 305 可以是液晶显示器（liquid crystal display，LCD），发光二极管（light emitting diode，LED）显示设备，阴极射线管（cathode ray tube，CRT）显示设备，或投影仪（projector）等。输入设备 306 和处理器 301 耦合，可以以多种方式接收用户的输入。例如，输入设备 306 可以是鼠标、键盘、触摸屏设备或传感设备等。

可以理解的，图 3 中示出的组成结构并不构成对该通信装置的限定，除图 3 所示部件之外，该通信装置可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件布置。

下面将结合附图，对本申请提供的方法进行描述。下述实施例中的各模块可以具备图 3 所示部件，不予赘述。

可以理解的是，本申请下述实施例中各个模块之间的消息名字或消息中各参数的名字等只是一个示例，具体实现中也可以是其他的名字，本申请对此不作具体限定。

可以理解的，本申请中“向……（如可信模块）发送第一请求信息”可以理解为该信息的目的端是可信模块，可以包括直接或间接的向可信模块发送信息。“接收来自……（如推理模块）的第一请求信息”可以理解为该信息的源端是推理模块，可以包括直接或间接的从推理模块接收信息。信息在信息发送的源端和目的端之间可能会被进行必要的处理，例如格式变化等，但目的端可以理解来自源端的有效信息。本申请中类似的表述可以做类似的理解，在此不再赘述。

可以理解的是，在本申请中，“/”可以表示前后关联的对象是一种“或”的关系，例如，A/B 可以表示 A 或 B；“和/或”可以用于描述关联对象存在三种关系，例如，A 和/或 B，可以表示：单独存在 A，同时存在 A 和 B，单独存在 B 这三种情况，其中 A，B 可以是单数或者复数。此外，类似于“A、B 和 C 中的至少一项”或“A、B 或 C 中的至少一项”的表述通常用于表示如下中任一项：单独存在 A；单独存在 B；单独存在 C；同时存在 A 和 B；同时存在 A 和 C；同时存在 B 和 C；同时存在 A、B 和 C。以上是以 A、B 和 C 共三个元素进行举例来说明该项目的可选用条目，当表述中具有更多元素时，该表述的含义可以按照前述规则获得。

为了便于描述本申请的技术方案，在本申请中，可以采用“第一”、“第二”等字样对功能相同或相似的技术特征进行区分。该“第一”、“第二”等字样并不对数量和执行次序进行限定，并且“第一”、“第二”等字样也并不限定一定不同。在本申请中，“示例性的”或者“例如”等词用于表示例子、例证或说明，被描述为“示例性的”或者“例如”的任何实施例或设计方案不应被解释为比其它实施例或设计方案更优选或更具优势。使用“示例性的”或者“例如”等词旨在以具体方式呈现相关概念，便于理解。

可以理解，说明书通篇中提到的“实施例”意味着与实施例有关的特定特征、结构或特性包括在本申请的至少一个实施例中。因此，在整个说明书各个实施例未必一定指相同的实施例。此外，这些特定

的特征、结构或特性可以任意适合的方式结合在一个或多个实施例中。可以理解，在本申请的各种实施例中，各过程的序号的大小并不意味着执行顺序的先后，各过程的执行顺序应以其功能和内在逻辑确定，而不应对本申请的实施过程构成任何限定。

可以理解，在本申请中，“用于指示”可以包括直接指示和间接指示，也可以包括显式指示和隐式指示。当描述某一指示信息用于指示 A 时，可以包括该指示信息直接指示 A 或间接指示 A，而并不代表该指示信息中一定携带有 A。将某一信息（如下文所述的第一指示信息等）所指示的信息称为待指示信息，则具体实现过程中，对待指示信息进行指示的方式有很多种，例如但不限于，可以直接指示待指示信息，如待指示信息本身或者该待指示信息的索引等。也可以通过指示其他信息来间接指示待指示信息，其中该其他信息与待指示信息之间存在关联关系。还可以仅仅指示待指示信息的一部分，而待指示信息的其他部分则是已知的或者提前约定的。例如，还可以借助预先约定（例如协议规定）的各个信息的排列顺序来实现对特定信息的指示，从而在一定程度上降低指示开销。

可以理解，在本申请中，“当……时”、“在……的情况下”、“若”以及“如果”均指在某种客观情况下会做出相应的处理，并非是限定时间，且也不要求实现时一定要有判断的动作，也不意味着存在其它限定。

可以理解，本申请中的一些可选的特征，在某些场景下，可以不依赖于其他特征，比如其当前所基于的方案，而独立实施，解决相应的技术问题，达到相应的效果，也可以在某些场景下，依据需求与其他特征进行结合。相应的，本申请中给出的装置也可以相应的实现这些特征或功能，在此不予赘述。

可以理解的，本申请中同一个步骤或者具有相同功能的步骤或者技术特征在不同实施例之间可以互相参考借鉴。

可以理解的，本申请中，可信模块，和/或，推理模块，和/或，管理模块可以执行本申请中的部分或全部步骤，这些步骤仅是示例，本申请还可以执行其它步骤或者各种步骤的变形。此外，各个步骤可以按照本申请呈现的不同的顺序来执行，并且有可能并非要执行本申请中的全部步骤。

可以理解的，本申请下述提供的方法中是以可信模块、推理模块和管理模块作为该交互示意的执行主体为例来示意该方法，但本申请并不限制该交互示意的执行主体。例如，本申请下述实施例提供的方法中的可信模块也可以是支持该可信模块实现该方法的芯片、芯片系统、或处理器，还可以是能实现全部或部分可信模块的逻辑节点、逻辑模块或软件；本申请下述提供的方法中的推理模块也可以是支持该推理模块实现该方法的芯片、芯片系统、或处理器，还可以是能实现全部或部分推理模块的逻辑节点、逻辑模块或软件；本申请下述提供的方法中的管理模块也可以是支持该管理模块实现该方法的芯片、芯片系统、或处理器，还可以是能实现全部或部分管理模块的逻辑节点、逻辑模块或软件。

如图 4 所示，为本申请提供的一种确定任务执行策略的方法，该方法可以包括如下步骤：

S401：推理模块获取第一任务的第一样本。

本申请中，推理模块可以是图 2A 所示的推理模块 202。第一任务为推理模块要执行的任务。第一任务可以对应至少一个样本，该至少一个样本包括第一样本。

一种可能的实现方式，推理模块从对应的 RAN 节点获取第一任务的第一样本。可以理解的，如果推理模块部署于该 RAN 节点，推理模块利用 RAN 节点内部的接口获取第一任务的第一样本。如果推理模块位于域管理功能（如 OAM），推理模块可以通过 OAM 与 RAN 节点的接口获取第一任务的第一样本。下面以推理模块部署于该 RAN 节点为例说明推理模块获取第一任务的第一样本的过程。

示例性的，以第一任务为波束管理用例为例进行说明。在全波束被配置为 64 波束、稀疏波束配置为 16 波束的情况下，RAN 节点通过向终端指示稀疏波束扫描来发起第一样本的采集，终端获取 64 波束中的 16 波束的波束测量结果（如 16 波束的 RSRP）作为稀疏波束扫描的结果，并将稀疏波束扫描的结果发送给 RAN 节点。RAN 节点接收到来自终端的稀疏波束扫描的结果后，将稀疏波束扫描的结果指示给推理模块，推理模块就可以将稀疏波束扫描的结果作为第一任务的第一样本。

示例性的，以第一任务为小区负载均衡用例为例进行说明。可以理解的，小区负载均衡是针对一个 RAN 节点管理的多个小区，如果这多个小区中可能有某些小区接入的终端数量将要达到或者已经超过小区的负载能力，会导致接入这些小区的终端的服务质量下降，而多个小区中的其他小区接入终端数量可能很少。在这种情况下，为了保证终端的服务质量，可以对小区之间进行终端的接入数量调整。推理模块可以指示 RAN 节点获取接入该 RAN 节点管理的多个小区的终端信息，终端信息可以包括终端的位置、移动性（mobility）以及一些测量指标的信息。其中，测量指标的信息可以包括如 RSRP，和/或，参考信号接收

质量（reference signal receiving quality，RSRQ），和/或，信号与干扰加噪声比（signal to interference plus noise ratio）等。推理模块从该RAN节点获取接入该RAN节点管理的多个小区的终端信息，可以将该终端信息作为第一任务的第一样本。

可选的，管理模块向推理模块发送第三指示信息。相应的，推理模块接收来自管理模块的第三指示信息。其中，第三指示信息可以指示第一任务。例如，第三指示信息可以包括第一任务的名称或标识。可选的，第三指示信息还可以包括第一任务对应的模型（如S402中的第一模型）的名称或标识。推理模块接收到第三指示信息后，可以根据第三指示信息获取第一任务的第一样本。其中，管理模块可以是图2A所示的管理模块203。

可以理解的，第一任务也可以预设置在推理模块中。如此，管理模块可以不需要向推理模块发送第三指示信息。

S402：推理模块向可信模块发送第一请求信息。相应的，可信模块接收来自推理模块的第一请求信息。

其中，第一请求信息用于指示第一任务对应的第一模型以及第一任务对应的第一样本。例如，第一请求信息包括第一模型的名称或第一模型的标识，以及第一样本。

本申请中，可信模块可以是图2A所示的可信模块201。

一种可能的实现方式，可信模块收到第一请求信息后，根据第一模型确定第一样本是否为对抗样本。具体的，可信模块可以在一定范围内对输入进行扰动，根据第一模型输出结果的变化程度确定第一样本是否为对抗样本。具体的，如果第一模型输出的变化量超过预设阈值，则可信模块确定第一样本处于模型的不稳定区域，将第一样本确定为对抗样本；如果该样本处于第一模型的稳定区域，可信模块确定第一样本确定为非对抗样本。

可选的，可信模块获取第一任务的鲁棒性需求，第一任务的鲁棒性需求指示需要检测第一任务对应的样本是否是对抗样本。这样，可信模块可以根据获取到的第一任务的鲁棒性需求，确定是否需要检测第一任务对应的样本是否是对抗样本。例如，管理模块向可信模块发送第一任务的鲁棒性需求。相应的，可信模块可以接收来自管理模块的第一任务的鲁棒性需求。应理解不同模型可以有不同的不稳定区域，所以同一任务的不同模型对应的鲁棒性需求可以相同或不同。不同任务对应的鲁棒性需求可以相同或不同。

可以理解的，第一任务的鲁棒性需求也可以指示不需要检测第一任务对应的第一样本是否是对抗样本，可信模块对第一样本不做是否为对抗样本的检测。

可选的，第一任务的鲁棒性需求还指示第一任务对应的第一模型的输出结果的鲁棒性要求，例如，第一任务的鲁棒性需求可以包括第一阈值。该第一阈值为对第一样本是否处于第一模型的不稳定区域判定的阈值。这样，可信模块可以根据第一模型和该鲁棒性要求确定第一样本是否为对抗样本。

一种可能的实现方式，第一任务对应的第一模型的输出结果的鲁棒性要求可以是通过上述第一阈值来量化。具体来说，上述可信模块对第一样本的输入进行扰动，根据第一模型输出的变化量是否大于或等于第一阈值，确定第一样本是否为对抗样本。可以理解的，鲁棒性要求越低，第一阈值越大，也就是说可信模块对第一样本判断为对抗样本的标准较低；鲁棒性要求越高，第一阈值越小，也就是说可信模块对第一样本判断为对抗样本的标准较高，第一样本越容易落在第一模型的不稳定区域内，相比于鲁棒性要求低的场景，第一样本越容易被判定为对抗样本。

S403:在第一样本为第一模型的对抗样本的情况下，可信模块向推理模块发送第一任务的执行策略。相应的，推理模块接收来自可信模块的第一任务的执行策略。

本申请中，第一任务的执行策略用于指示将第一样本变更为非对抗样本的方式。例如，该方式包括：更换第一模型；或者，更换第一样本；或者，对第一样本执行第一操作。其中，第一操作包括以下至少一项：特征压缩、样本去噪或数据平滑处理。

一种可能的实现方式，可信模块获取第一指示信息。其中，第一指示信息用于指示第一任务是否有备用样本，和/或，第一指示信息用于指示第一任务是否有备用模型。例如，推理模块向可信模块发送第一指示信息。相应的，可信模块接收来自推理模块的第一指示信息。如此，可信模块可以根据第一指示信息确定第一任务的执行策略。可选的，在推理模块向可信模块发送第一指示信息之前，管理模块可以向推理模块发送第四指示信息。其中，第四指示信息可以指示第一任务是否有备用模型。这样，推理模块接收到第四指示信息后可以向可信模块发送第一指示信息。

一种可能的设计，第一指示信息可以包括备用模型的 ID，以向可信模块指示第一任务有备用模型。具体的，备用模型可以是第一模型的不同版本，或者备用模型与第一模型的输入或输出类似，但是备用模型与第一模型的内部实现可以不尽相同。例如，以负载均衡用例的模型为例，说明备用模型和第一模型的关系：备用模型的输入与第一模型的输入（如第一样本）可以均包括RAN节点资源使用情况、终端性能数据、邻区节点资源使用情况和对应的终端性能，但是备用模型进行负载均衡的算法与第一模型不同。例如，RAN 节点可以向除该 RAN 节点之外的其他 RAN 节点获取其他 RAN 节点部署的负载均衡用例的模型，将获取到的模型作为第一任务的备用模型。

可以理解的，第一指示信息指示的内容不同，第一任务的执行策略指示的将第一样本变更为非对抗样本的方式不同。下面进行具体阐述。

设计 1，若第一指示信息指示第一任务有备用样本，则第一任务的执行策略指示更换第一样本。换言之，若第一指示信息指示第一任务有备用样本，则可信模块可以指示推理模块通过更换第一样本的方式将第一样本变更为非对抗样本。

设计 2，若第一指示信息指示第一任务有备用模型，则第一任务的执行策略指示更换第一模型。换言之，若第一指示信息指示第一任务有备用模型，则可信模块可以指示推理模块通过更换第一模型的方式将第一样本变更为非对抗样本。

可以理解的，本申请不限制备用模型的数量。当第一任务的备用模型有多个时，第一任务的执行策略可以指示该多个备用模型中的全部或部分模型，以表示可以通过这些备用模型进行模型推理。可选的，第一任务的执行策略还可以包括：合并多个备用模型的推理结果的方法，例如投票、加权平均等方法。例如，以第一任务为分类任务为例说明投票方法：投票可以是指根据多个备用模型对第一样本分别进行样本推理后输出的结果的置信度，选择置信度最高的结果作为第一样本的分类结果。又例如，以第一任务为回归任务为例说明加权平均方法：加权平均可以是指根据多个备用模型对第一样本分别进行样本推理后输出的结果的数值，分别进行加权平均后的数值作为第一任务的输出结果。其中，每个备用模型输出的结果对应的权重可以是预设置的，或者是相同的，本申请不做限定。

一种可能的实现方式，第一任务的执行策略指示更换第一模型时，推理模块接收来自通信节点的第二指示信息，其中，第二指示信息用于指示第一任务的备用模型。其中，通信节点可以是与部署第一模型的RAN节点相邻的RAN节点等，不作限制。例如，第二指示信息包括第一任务的备用模型。又例如，第二指示信息包括第一任务的备用模型的ID，推理模块可以根据备用模型的ID向管理模块获取备用模型的ID对应的备用模型。

可选的，推理模块可以向通信节点请求第一任务的备用模型，通信节点基于推理模块的请求向推理模块发送第二指示信息。相应的，推理模块接收来自通信节点的第二指示信息。

另一种可能的实现方式，推理模块向通信节点发送第一样本。该通信节点存储有第一任务的备用模型。该通信节点接收到第一样本后，可以将第一样本输入第一任务的备用模型，得到推理结果，并向推理模块发送推理结果。相应的，推理模块接收来自通信节点的推理结果。其中，推理结果用于指示根据第一样本和第一任务的备用模型进行推理得到的结果。示例性的，该通信节点为与部署第一模型的RAN节点相邻的RAN节点的。

可以理解的，第一任务可以对应一个或多个备用模型。当第一任务对应一个备用模型或多个备用模型，并且这些备用模型部署在一个通信节点时，推理模块向这一个通信节点发送第一样本，以请求该通信节点根据第一样本得到该通信节点部署的备用模型的推理结果。当第一任务对应多个备用模型，并且该多个备用模型部署在不同的通信节点时，推理模块还可以分别向上述不同的通信节点发送第一样本，以请求上述不同的通信节点中的每个通信节点分别根据第一样本得到每个通信节点部署的备用模型的推理结果。

设计 3，若第一指示信息指示第一任务没有备用样本也没有备用模型，将第一样本变更为非对抗样本的方式包括对第一样本执行第一操作。换言之，若第一指示信息指示第一任务没有备用样本也没有备用模型，则可信模块可以通过对第一样本执行第一操作将第一样本变更为非对抗样本。

应理解，如果第一样本被确定为第一模型的对抗样本，利用第一模型对第一样本进行样本推理很可能会导致第一模型推理出错误的结果，为了避免出现这种情况，可以用其他样本进行样本推理。在无备用样本的情况下，一种可能的实现方式是，可信模块可以对第一样本执行第一操作，以剔除导致第一模型推理出错的样本数据。例如，可信模块可以通过对第一样本进行特征压缩、样本去噪或数据平滑等方

法，剔除导致第一模型推理出错的样本数据，使执行第一操作之后得到的样本变更为第一模型的非对抗样本。

可选的，第一操作的具体方法（即特征压缩、样本去噪或数据平滑等方法）可以由管理模块进行指示。如果可信模块未收到管理模块指示第一操作的具体方法，可信模块可以根据第一样本的特征选择合适第一操作，或者第一任务为回归任务时，第一操作可以是样本去噪或数据平滑等，不作限制。

其中，样本去噪即去除样本中的噪声，噪声是第一样本中存在的错误或异常值，去除噪声可以避免噪声对第一模型的误导。特征压缩可以是通过选取数据中的分类信息或判别特征，去除冗余信息的方法。数据平滑可以对数据抖动剧烈的第一样本进行处理，得到数值较为稳定的数据样本。

可以理解的，上述设计 1~设计 3 仅是第一指示信息与第一任务的执行策略的对应关系的示例，在具体应用中，第一指示信息与第一任务的执行策略还可以有其他的对应关系。例如，在设计 1 或设计 2 中，将第一样本变更为非对抗样本的方式还可以替换为对第一样本执行第一操作。

可选的，当第一任务的执行策略指示多项时，将执行策略的优先级按照从高到低的顺序可以排列为：更换第一样本，更换第一模型以及对第一样本执行第一操作。

除了上述方式之外，可信模块可以根据第一策略信息确定第一任务的执行策略。例如，管理模块向可信模块发送第一策略信息，相应的，可信模块接收来自管理模块的第一策略信息。第一策略信息指示更换第一模型；或者，更换第一样本；或者，对第一样本执行第一操作。

示例性的，管理模块通过第一策略信息向可信模块指示第一任务的执行策略指示更换第一模型时，可信模块接收到来自推理模块的第一指示信息也指示了第一任务有备用模型，可信模块可以将第一任务的执行策略确定为更换第一模型。

可选的，可信模块确定第一任务的执行策略后，可以向推理模块发送该策略，以便推理模块根据该策略执行相应的操作。可选的，第一任务的执行策略还指示第一样本为对抗样本。

可选的，第一任务的执行策略还指示对第一样本执行第一操作得到的样本。若可信模块对第一样本执行第一操作可以得到多个样本，第一任务的执行策略还可以指示第一模型对该多个样本分别进行推理得到的结果的合并方法，合并方法可以是投票、加权平均等，其中，投票或加权平均的具体方法可以参见设计 2 中对投票或加权平均过程的解释，不再赘述。

可以理解的，当可信模块将第一样本确定为非对抗样本时，也可以通过第一任务的执行策略进行指示。例如第一任务的执行策略可以为空的消息，以便推理模块获知第一样本为非对抗样本。当可信模块将第一样本确定为非对抗样本时，推理模块可以根据第一任务的执行策略获知第一样本为对抗样本，以及将对抗样本变更为非对抗样本的方式。

可以理解的，若第一任务的执行策略指示更换第一模型，则推理模块可以将第一样本输入备用模型，进行模型推理。若第一任务的执行策略包括合并多个备用模型的推理结果的方法，则推理模块可以根据合并多个备用模型的推理结果的方法，确定出多个备选模型的推理结果，然后根据多个备选模型的推理结果确定出最终输出结果。示例性的，以第一任务为负载均衡用例为例进行说明。若第一任务的执行策略指示第一样本为对抗样本、备用模型包括备用模型 1 和备用模型 2，合并备用模型 1 和备用模型 2 的推理结果的方法为投票，则推理模块可以将第一样本分别输入备用模型 1 和备用模型 2 进行模型推理，得到备用模型 1 的推理结果和备用模型 2 的推理结果，并将二者中终端调度最少的小区负载均衡方案作为最终结果。或者，若第一任务的执行策略指示更换第一模型，则推理模块可以获取第一任务对应的第二模型，并请求可信模块确定第一样本是否是第二模型的对抗样本，具体的，可以参考下述 S407~S408 中对应的描述，不做赘述。

可以理解的，若第一任务的执行策略指示更换第一样本，则推理模块可以获取第一任务对应的第二样本，并请求可信模块确定第二样本是否是第一模型对抗样本，具体的，可以参考下述 S405~S406 中对应的描述，不做赘述。

可以理解的，若第一任务的执行策略指示对第一样本执行第一操作、以及对第一样本执行第一操作得到的样本，推理模块可以将接收到的样本输入第一模型进行模型推理。若对第一样本执行第一操作得到多个样本，第一任务的执行策略还指示该多个样本对应的多样本推理结果合并方法，则推理模块可以将每个样本输入第一模型，并根据该方式对多个样本推理得到的结果进行合并。

基于图 4 所示的方法，如果第一任务对应的第一样本为第一任务对应的第一模型的对抗样本，可信模块根据第一指示信息确定出第一任务的执行策略，由此可以将第一样本变更为非对抗样本，使第一任务

对应的模型能够输出正确的结果，从而降低模型推理的出错率。

可选的，在图 4 所示方法的一种可能的实现方式中，若第一任务的执行策略指示更换第一样本，可信模块还可以获取第一任务对应的第二样本，并根据第一模型确定第二样本是否是对抗样本，以便降低第一模型推理的出错率。具体的可以如图 5 所示，图 4 所示的方法还包括如下步骤：

S404：可信模块获取第一任务对应的第二样本。

一种可能的实现方式，当第一样本被可信模块判定为第一模型的对抗样本时，可信模块可以通过推理模块获取第一任务对应的第二样本。第二样本为第一任务的备用样本。

示例性的，以第一任务为波束管理用例为例说明，推理模块根据第一任务的执行策略获知第一样本是第一模型的对抗样本，以及更换第一样本后，指示 RAN 节点收集终端侧稀疏波束扫描结果，RAN 节点将终端反馈的一组稀疏波束扫描结果（即第二样本）指示给推理模块。后续，推理模块可以向可信模块发送该扫描结果。

S405：可信模块根据第一模型确定第二样本是否是对抗样本。

可信模块判定第二样本是否是第一模型的对抗样本。如果第二样本是第一模型的非对抗样本，可信模块可以将该结果指示给推理模块，推理模块可根据第二模型对第一样本进行推理，得到推理结果。如果是第二样本是第一模型的对抗样本，可信模块可以重新确定第一任务的执行策略（例如，该策略为更换其他备用样本），并将该策略发送给推理模块，以便降低第一模型的出错率。具体过程可以参考 S403，不再赘述。

可选的，在图 4 所示方法的一种可能的实现方式中，若第一任务的执行策略指示更换第一模型，可信模块还可以获取第一任务对应的第二模型，并根据第二模型确定第一样本是否是对抗样本，以便降低第一模型推理的出错率。具体的可以如图 5 所示，图 4 所示的方法还包括如下步骤：

S406：可信模块获取第一任务对应的第二模型。

一种可能的实现方式，可信模块可以通过推理模块获取第一任务对应的第二模型，推理模块可以通过管理模块在第四指示信息中指示的备选模型作为第二模型。

S407：可信模块根据第二模型确定第一样本是否是对抗样本。

可信模块判定第一样本是否是第二模型的对抗样本的具体过程可以参考 S403，不再赘述。

一种可能的实现方式，如果可信模块确认第一样本是第二模型的非对抗样本，可以将该结果指示给推理模块，推理模块可根据第二模型对第一样本进行推理，得到推理结果。

可以理解的，如果可信模块确认第一样本是第二模型的对抗样本，并将该结果指示给推理模块，推理模块如果能获取到第一任务的其他备用模型，可以继续让可信模块确认第一样本是否是第一任务的其他备用模型的对抗样本。可选的，可信模块也可以指示推理模块对第一样本执行第一操作；或者在有第二模型的备用样本的情况下，可信模块可以指示更换第一样本，不做限定。

上述主要从可信模块与推理模块之间交互的角度对本申请提供的方案进行了介绍。相应的，本申请还提供了通信装置，该通信装置可以为上述方法实施例中的可信模块，或者包含上述可信模块的装置，或者为可用于可信模块的部件；或者，该通信装置可以为上述方法实施例中的推理模块，或者包含上述推理模块的装置，或者为可用于推理模块的部件。可以理解的是，上述可信模块或推理模块等为了实现上述功能，其包含了执行各个功能相应的硬件结构和/或软件模块。本领域技术人员应该很容易意识到，结合本文中所公开的实施例描述的各示例的单元及算法操作，本申请能够以硬件或硬件和计算机软件的结合形式来实现。某个功能究竟以硬件还是计算机软件驱动硬件的方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。

本申请可以根据上述方法示例对可信模块和推理模块进行功能模块的划分，例如，可以对应各个功能划分各个功能模块，也可以将两个或两个以上的功能集成在一个处理模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。可以理解的是，本申请中对模块的划分是示意性的，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式。

比如，以采用集成的方式划分各个功能模块的情况下，图 6 示出了一种通信装置 60 的结构示意图。通信装置 60 包括接口模块 601 和处理模块 602。接口模块 601，也可以称为接口单元，用于执行收发操作，例如可以是接口电路，收发机，收发器或者通信接口等。处理模块 602，也可以称为处理单元，用于执行除了收发操作之外的操作，例如可以是处理电路或者处理器等。

在一些实施例中，该通信装置60还可以包括存储模块（图6中未示出），用于存储程序指令和数据。

示例性地，通信装置60用于实现可信模块的功能。通信装置60例如为图4所示的实施例或图5所示的实施例的可信模块。

其中，接口模块601，用于接收第一请求信息。其中，第一请求信息用于指示第一任务对应的第一模型以及第一任务对应的第一样本。例如，接口模块601可以用于执行S402。

处理模块602，用于在第一样本为第一模型的对抗样本的情况下，控制接口模块601发送第一任务的执行策略。其中，第一任务的执行策略用于指示将第一样本变更为非对抗样本的方式。例如，处理模块602可以用于执行S403。

在一种可能的实现方式中，第一样本变更为非对抗样本的方式包括：更换第一模型；或者，更换第一样本；或者，对第一样本执行第一操作。

在一种可能的实现方式中，处理模块602，还用于获取第一指示信息，第一指示信息用于指示第一任务是否有备用样本，和/或，第一指示信息用于指示第一任务是否有备用模型；根据第一指示信息确定第一任务的执行策略。

在一种可能的实现方式中，第一指示信息指示第一任务有备用样本，第一任务的执行策略指示更换第一样本；或者，第一指示信息指示第一任务有备用模型，第一任务的执行策略指示更换第一模型；第一指示信息指示第一任务没有备用样本也没有备用模型，第一任务的执行策略指示对第一样本执行第一操作。

在一种可能的实现方式中，第一操作包括以下至少一项：特征压缩、样本去噪或数据平滑处理。

在一种可能的实现方式中，第一任务的执行策略指示更换第一样本，处理模块602，还用于获取第一任务对应的第二样本；根据第一模型确定第二样本是否是对抗样本。

在一种可能的实现方式中，第一任务的执行策略指示更换第一模型，处理模块602，还用于获取第一任务对应的第二模型；根据第二模型确定第一样本是否是对抗样本。

在一种可能的实现方式中，处理模块602，还用于接收第一策略信息，第一策略信息指示第一任务的执行策略。

在一种可能的实现方式中，处理模块602，还用于获取第一任务的鲁棒性需求，第一任务的鲁棒性需求指示需要检测第一任务对应的第一样本是否是对抗样本；处理模块602，还用于根据第一模型确定第一样本是否为对抗样本。

在一种可能的实现方式中，第一任务的鲁棒性需求还指示第一任务对应的第一模型的输出结果的鲁棒性要求；处理模块602，具体用于根据第一模型和鲁棒性要求确定第一样本是否为对抗样本。

当用于实现可信模块的功能时，关于通信装置60所能实现的其他功能，可参考图4所示的实施例或图5所示的实施例的相关介绍，不多赘述。

或者，示例性地，通信装置60用于实现推理模块的功能。通信装置60例如为图4所示的实施例，图5所示的实施例的推理模块。

其中，处理模块602，用于获取第一任务的第一样本。例如，处理模块602可以用于执行S401。

接口模块601，用于发送第一请求信息。其中，第一请求信息指示第一任务的第一样本和第一任务的第一模型。例如，接口模块601可以用于执行S402。

接口模块601，还用于接收第一任务的执行策略，第一任务的执行策略指示用于将第一样本变更为非对抗样本的方式。例如，接口模块601可以用于执行S403。

在一种可能的实现方式中，将第一样本变更为非对抗样本的方式包括：更换第一模型；或者，更换第一样本；或者，对第一样本执行第一操作。

在一种可能的实现方式中，接口模块601，还用于发送第一指示信息，第一指示信息用于指示第一任务是否有备用样本，和/或，第一指示信息用于指示第一任务是否有备用模型，第一指示信息用于确定第一任务的执行策略。

在一种可能的实现方式中，第一指示信息指示第一任务有备用样本，第一任务的执行策略指示更换第一样本；或者，第一指示信息指示第一任务有备用模型，第一任务的执行策略指示更换第一模型；第一指示信息指示第一任务没有备用样本也没有备用模型，第一任务的执行策略指示对第一样本执行第一操作。

在一种可能的实现方式中，第一操作包括以下至少一项：特征压缩、样本去噪或数据平滑处理。

在一种可能的实现方式中，第一任务的执行策略指示更换第一模型，接口模块 601，还用于接收来自通信节点的第二指示信息，第二指示信息用于指示第一任务的备用模型。

在一种可能的实现方式中，第一任务的执行策略指示更换第一模型，接口模块 601，还用于向通信节点发送第一样本；接口模块 601，还用于接收来自通信节点的推理结果，推理结果用于指示根据第一样本和第一任务的备用模型进行推理得到的结果。

在一种可能的实现方式中，接口模块 601，还用于接收第一任务的鲁棒性需求，第一任务的鲁棒性需求指示需要检测第一任务对应的样本是否是对抗样本。

在一种可能的实现方式中，第一任务的鲁棒性需求还指示第一任务的模型的输出结果的鲁棒性要求。

在一种可能的实现方式中，第一任务的执行策略还用于指示第一样本为对抗样本。

当用于推理模块的功能时，关于通信装置 60 所能实现的其他功能，可参考图 4 或图 5 所示的实施例的相关介绍，不多赘述。

在一个简单的实施例中，本领域的技术人员可以想到通信装置 60 可以采用图 3 所示的形式。比如，图 3 中的处理器 301 可以通过调用存储器 303 中存储的计算机执行指令，使得通信装置 60 执行上述实施例中所述的方法。

示例性的，图 6 中的接口模块 601 和处理模块 602 的功能/实现过程可以通过图 3 中的处理器 301 调用存储器 303 中存储的计算机执行指令来实现。或者，图 6 中的处理模块 602 的功能/实现过程可以通过图 3 中的处理器 301 调用存储器 303 中存储的计算机执行指令来实现，图 6 中的接口模块 601 的功能/实现过程可以通过图 3 中的通信接口 304 来实现。

可以理解的是，以上模块或单元的一个或多个可以软件、硬件或二者结合来实现。当以上任一模块或单元以软件实现的时候，所述软件以计算机程序指令的方式存在，并被存储在存储器中，处理器可以用于执行所述程序指令并实现以上方法流程。该处理器可以内置于 SoC（片上系统）或 ASIC，也可是一个独立的半导体芯片。该处理器内处理用于执行软件指令以进行运算或处理的核外，还可进一步包括必要的硬件加速器，如现场可编程门阵列（field programmable gate array，FPGA）、PLD（可编程逻辑器件）、或者实现专用逻辑运算的逻辑电路。

当以上模块或单元以硬件实现的时候，该硬件可以是 CPU、微处理器、数字信号处理（digital signal processing，DSP）芯片、微控制单元（microcontroller unit，MCU）、人工智能处理器、ASIC、SoC、FPGA、PLD、专用数字电路、硬件加速器或非集成的分立器件中的任一个或任一组合，其可以运行必要的软件或不依赖于软件以执行以上方法流程。

可选的，本申请还提供了一种芯片系统，包括：至少一个处理器和接口，该至少一个处理器通过接口与存储器耦合，当该至少一个处理器执行存储器中的计算机程序或指令时，使得上述任一方法实施例中的方法被执行。在一种可能的实现方式中，该芯片系统还包括存储器。可选的，该芯片系统可以由芯片构成，也可以包含芯片和其他分立器件，本申请对此不作具体限定。

可选的，本申请还提供了一种计算机可读存储介质。上述方法实施例中的全部或者部分流程可以由计算机程序来指令相关的硬件完成，该程序可存储于上述计算机可读存储介质中，该程序在执行时，可包括如上述各方法实施例的流程。计算机可读存储介质可以是前述任一实施例的通信装置的内部存储单元，例如通信装置的硬盘或内存。上述计算机可读存储介质也可以是上述通信装置的外部存储设备，例如上述通信装置上配备的插接式硬盘，智能存储卡（smart media card，SMC），安全数字（secure digital，SD）卡，闪存卡（flash card）等。进一步地，上述计算机可读存储介质还可以既包括上述通信装置的内部存储单元也包括外部存储设备。上述计算机可读存储介质用于存储上述计算机程序以及上述通信装置所需的其他程序和数据。上述计算机可读存储介质还可以用于暂时地存储已经输出或者将要输出的数据。

可选的，本申请还提供了一种计算机程序产品。上述方法实施例中的全部或者部分流程可以由计算机程序来指令相关的硬件完成，该程序可存储于上述计算机程序产品中，该程序在执行时，可包括如上述各方法实施例的流程。

可选的，本申请还提供了一种计算机指令。上述方法实施例中的全部或者部分流程可以由计算机指令来指令相关的硬件（如计算机、处理器、可信模块或推理模块等）完成。该程序可被存储于上述计算机可读存储介质中或上述计算机程序产品中。

可选的，本申请还提供了一种通信系统，包括：上述实施例中的可信模块和推理模块。可选的，该通信系统还包括：管理模块。

通过以上的实施方式的描述，所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，仅以上述各功能模块的划分进行举例说明，实际应用中，可以根据需要而将上述功能分配由不同的功能模块完成，即将装置的内部结构划分成不同的功能模块，以完成以上描述的全部或者部分功能。

在本申请所提供的几个实施例中，应该理解到，所揭露的装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述模块或单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个装置，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是一个物理单元或多个物理单元，即可以位于一个地方，或者也可以分布到多个不同地方。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

以上所述，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，任何在本申请揭露的技术范围内的变化或替换，都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应以所述权利要求的保护范围为准。

权 利 要 求 书

1．一种确定任务执行策略的方法，其特征在于，所述方法包括：

接收第一请求信息；所述第一请求信息用于指示第一任务对应的第一模型以及所述第一任务对应的第一样本；

在所述第一样本为所述第一模型的对抗样本的情况下，发送所述第一任务的执行策略，所述第一任务的执行策略用于指示将所述第一样本变更为非对抗样本的方式。

2．根据权利要求 1 所述的方法，其特征在于，所述将所述第一样本变更为非对抗样本的方式包括：

更换所述第一模型；或者，

更换所述第一样本；或者，

对所述第一样本执行第一操作。

3．根据权利要求 1 或 2 所述的方法，其特征在于，所述方法还包括：

获取第一指示信息，所述第一指示信息用于指示所述第一任务是否有备用样本，和/或，所述第一指示信息用于指示所述第一任务是否有备用模型；

根据所述第一指示信息确定所述第一任务的执行策略。

4．根据权利要求 3 所述的方法，其特征在于，所述根据所述第一指示信息确定所述第一任务的执行策略，包括：

所述第一指示信息指示所述第一任务有备用样本，所述第一任务的执行策略指示更换所述第一样本；或者，

所述第一指示信息指示所述第一任务有备用模型，所述第一任务的执行策略指示更换所述第一模型；

所述第一指示信息指示所述第一任务没有备用样本也没有备用模型，所述第一任务的执行策略指示对所述第一样本执行第一操作。

5．根据权利要求 2 或 4 所述的方法，其特征在于，所述第一操作包括以下至少一项：特征压缩、样本去噪或数据平滑处理。

6．根据权利要求 2、4 或 5 所述的方法，其特征在于，所述第一任务的执行策略指示更换所述第一样本，所述方法还包括：

获取所述第一任务对应的第二样本；

根据所述第一模型确定所述第二样本是否是对抗样本。

7．根据权利要求 2、4 或 5 所述的方法，其特征在于，所述第一任务的执行策略指示更换所述第一模型，所述方法还包括：

获取所述第一任务对应的第二模型；

根据所述第二模型确定所述第一样本是否是对抗样本。

8．根据权利要求 1 或 2 所述的方法，其特征在于，所述方法还包括：

接收第一策略信息，所述第一策略信息指示所述第一任务的执行策略。

9．根据权利要求 1-8 中任一项所述的方法，其特征在于，所述方法还包括：

获取所述第一任务的鲁棒性需求，所述第一任务的鲁棒性需求指示需要检测所述第一任务对应的所述第一样本是否是对抗样本；

根据所述第一模型确定所述第一样本是否为对抗样本。

10．根据权利要求 9 所述的方法，其特征在于，所述第一任务的鲁棒性需求还指示所述第一任务对应的所述第一模型的输出结果的鲁棒性要求；

所述根据所述第一模型确定所述第一样本是否为对抗样本，包括：

根据所述第一模型和所述鲁棒性要求确定所述第一样本是否为对抗样本。

11.一种确定任务执行策略的方法，其特征在于，所述方法包括：

获取第一任务的第一样本；

发送第一请求信息；所述第一请求信息指示所述第一任务的第一样本和所述第一任务的第一模型；

接收第一任务的执行策略，所述第一任务的执行策略指示用于将所述第一样本变更为非对抗样本的方式。

12．根据权利要求 11 所述的方法，其特征在于，所述将所述第一样本变更为非对抗样本的方式包括：

更换所述第一模型；或者，

更换所述第一样本；或者，

对所述第一样本执行第一操作。

13．根据权利要求 11 或 12 所述的方法，其特征在于，所述方法还包括：

发送第一指示信息，所述第一指示信息用于指示所述第一任务是否有备用样本，和/或，所述第一指示信息用于指示所述第一任务是否有备用模型，所述第一指示信息用于确定所述第一任务的执行策略。

14．根据权利要求 13 所述的方法，其特征在于，

所述第一指示信息指示所述第一任务有备用样本，所述第一任务的执行策略指示更换所述第一样本；或者，

所述第一指示信息指示所述第一任务有备用模型，所述第一任务的执行策略指示更换所述第一模型；

所述第一指示信息指示所述第一任务没有备用样本也没有备用模型，所述第一任务的执行策略指示对所述第一样本执行第一操作。

15．根据权利要求 12 或 14 所述的方法，其特征在于，所述第一操作包括以下至少一项：特征压缩、样本去噪或数据平滑处理。

16．根据权利要求 12-15 中任一项所述的方法，其特征在于，所述第一任务的执行策略指示更换所述第一模型，所述方法还包括：

接收来自通信节点的第二指示信息，所述第二指示信息用于指示所述第一任务的备用模型。

17．根据权利要求 12-16 中任一项所述的方法，其特征在于，所述第一任务的执行策略指示更换所述第一模型，所述方法还包括：

向通信节点发送所述第一样本；

接收来自所述通信节点的推理结果，所述推理结果用于指示根据所述第一样本和所述第一任务的备用模型进行推理得到的结果。

18．根据权利要求 11-17 中任一项所述的方法，其特征在于，所述方法还包括：

接收所述第一任务的鲁棒性需求，所述第一任务的鲁棒性需求指示需要检测所述第一任务对应的样本是否是对抗样本。

19．根据权利要求 18 所述的方法，其特征在于，所述第一任务的鲁棒性需求还指示所述第一任务的模型的输出结果的鲁棒性要求。

20．根据权利要求 11-19 中任一项所述的方法，其特征在于，所述第一任务的执行策略还用于指示所述第一样本为对抗样本。

21．一种通信装置，其特征在于，包括用于执行如权利要求 1 至 10 中任一项所述方法的单元或模块，或者包括用于执行如权利要求 11 至 20 中任一项所述方法的单元或模块。

22．一种通信装置，其特征在于，包括：处理器，所述处理器与存储器耦合，所述存储器用于存储程序或指令，当所述程序或指令被所述处理器执行时，使得所述装置执行如权利要求 1 至 10 中任一项所述的方法，或者执行如权利要求 11 至 20 中任一项所述的方法。

23．一种计算机可读存储介质，其上存储有计算机程序或指令，其特征在于，所述计算机程序或指令被执行时使得计算机执行如权利要求 1 至 10 中任一项所述的方法或者如权利要求 11 至 20 中任一项所述的方法。

24．一种计算机程序产品，所述计算机程序产品中包括计算机程序代码，其特征在于，当所述计算机程序代码在计算机上运行时，使得计算机实现权利要求 1 至 10 中任一项所述的方法或者实现权利要求 11 至 20 中任一项所述的方法。

25．一种通信系统，其特征在于，包括：用于执行如权利要求 1 至 10 中任一项所述方法的装置和用于执行如权利要求 11 至 20 中任一项所述方法的装置。

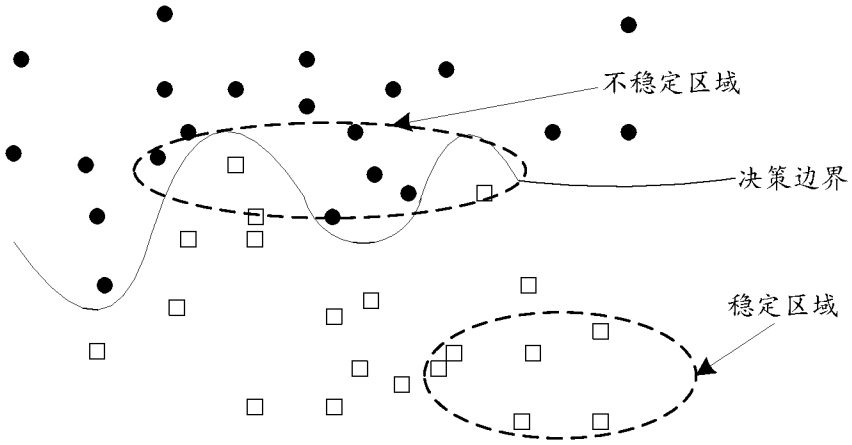
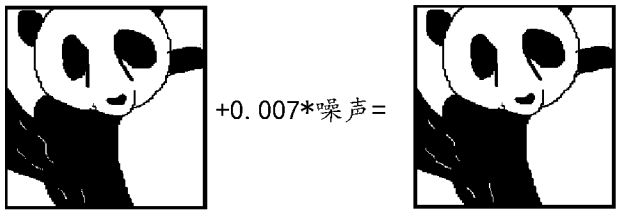


图 1A



原图片 对抗样本图片

图 1B

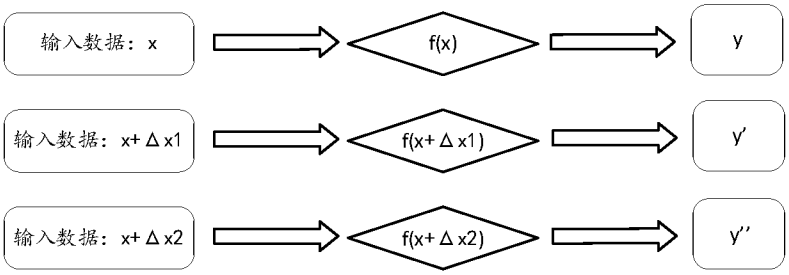


图 1C

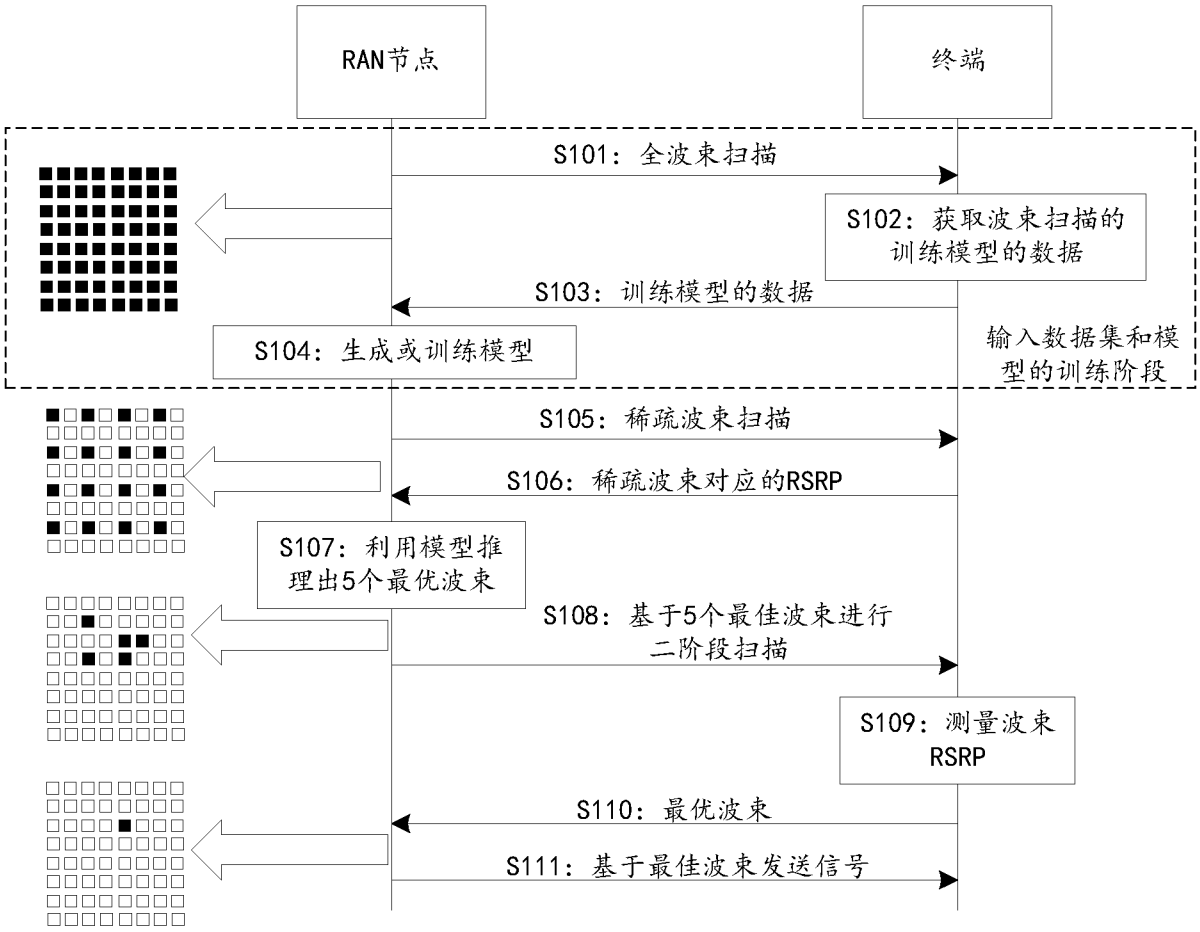

RAN节点
终端
S101：全波束扫描
S102：获取波束扫描的训练模型的数据
S103：训练模型的数据
S104：生成或训练模型
输入数据集和模型的训练阶段
S105：稀疏波束扫描
S106：稀疏波束对应的RSRP
S107：利用模型推理出5个最优波束
S108：基于5个最佳波束进行二阶段扫描
S109：测量波束RSRP
S110：最优波束
S111：基于最佳波束发送信号

图 1D

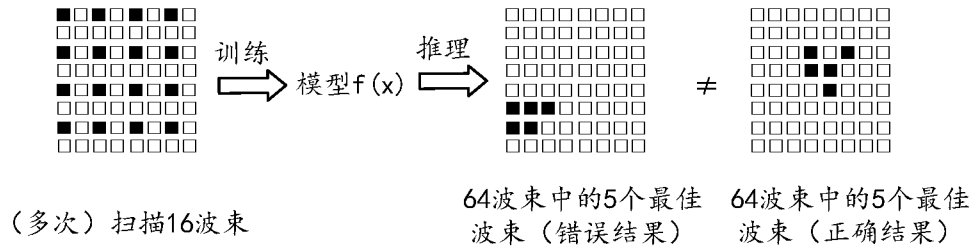

训练
模型f(x)
推理
≠
（多次）扫描16波束
64波束中的5个最佳波束（错误结果）
64波束中的5个最佳波束（正确结果）

图 1E

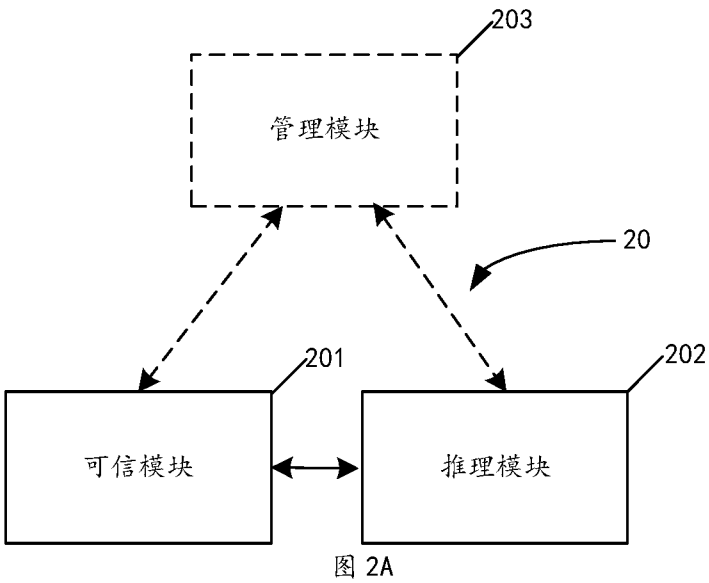

203
管理模块
20
201
202
可信模块
推理模块

图 2A

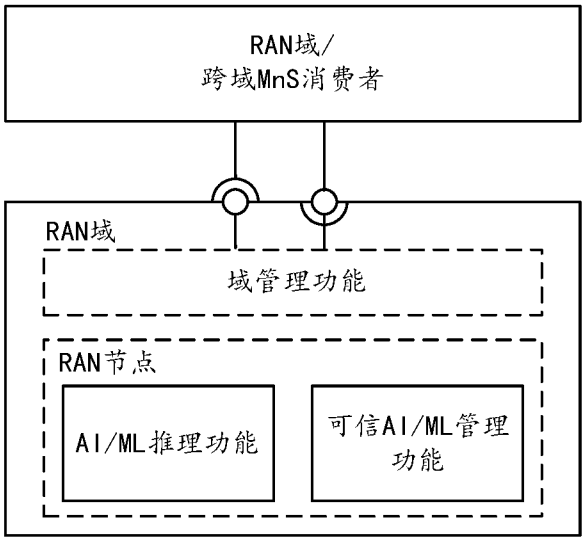

RAN域/
跨域MnS消费者
RAN域
域管理功能
RAN节点
AI/ML推理功能
可信AI/ML管理功能

图 2B

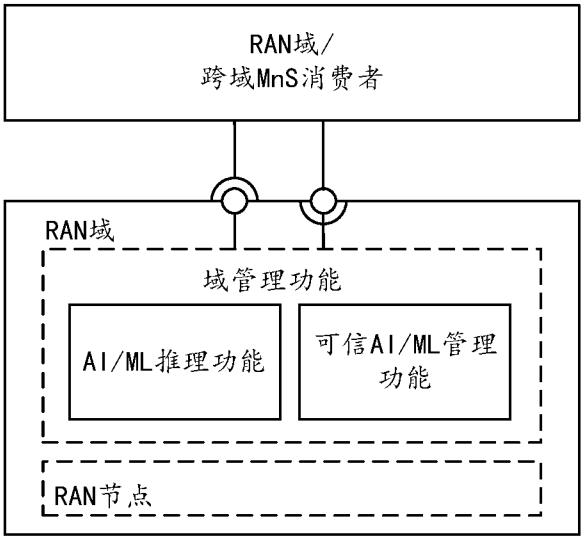

RAN域/
跨域MnS消费者
RAN域
域管理功能
AI/ML推理功能
可信AI/ML管理功能
RAN节点

图 2C

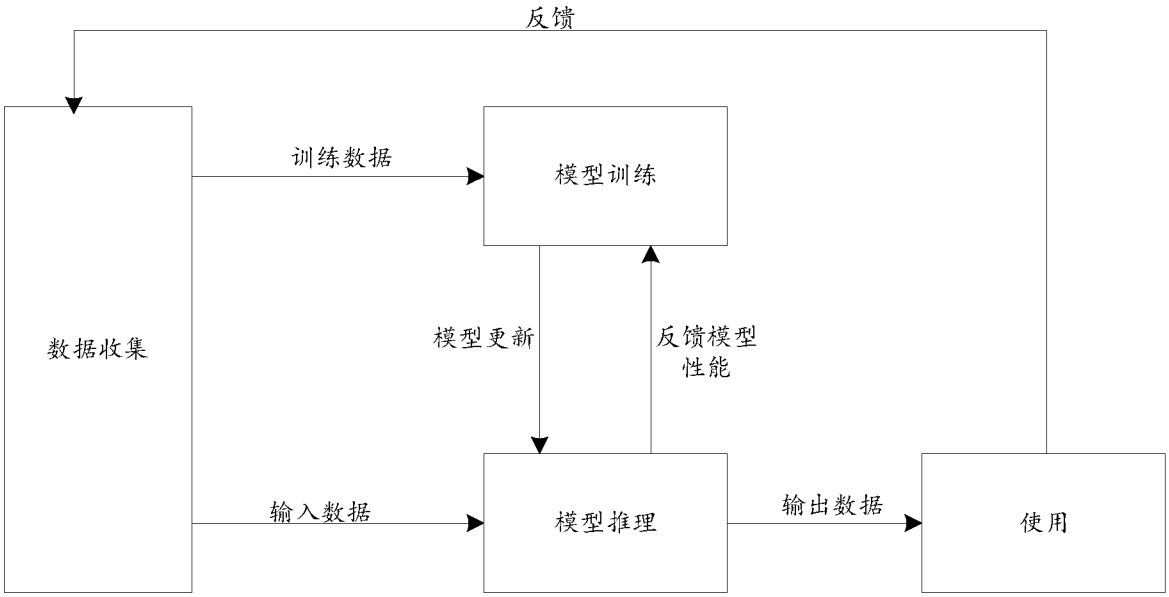

反馈
数据收集
训练数据
模型训练
模型更新
反馈模型性能
输入数据
模型推理
输出数据
使用

图 2D

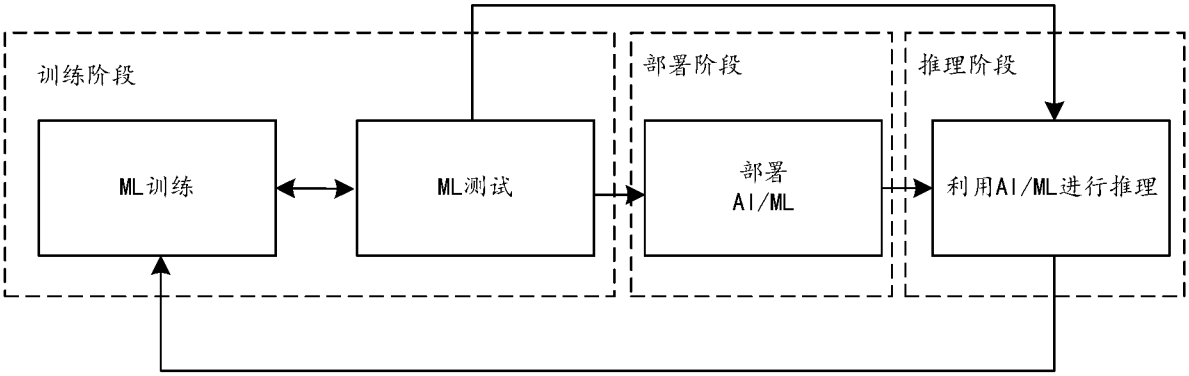

训练阶段
部署阶段
推理阶段
ML训练
ML测试
部署
AI/ML
利用AI/ML进行推理

图 2E

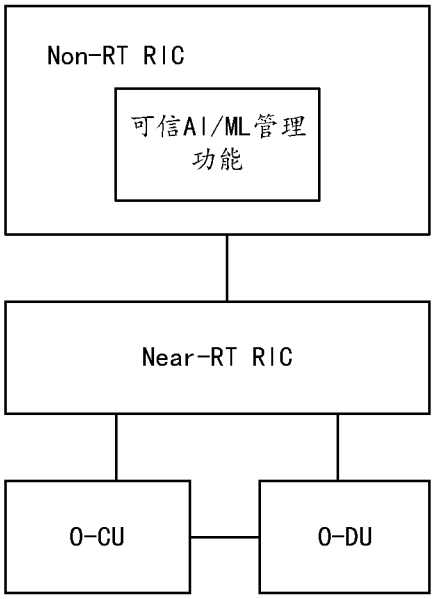

Non-RT RIC
可信AI/ML管理
功能
Near-RT RIC
O-CU
O-DU

图 2F

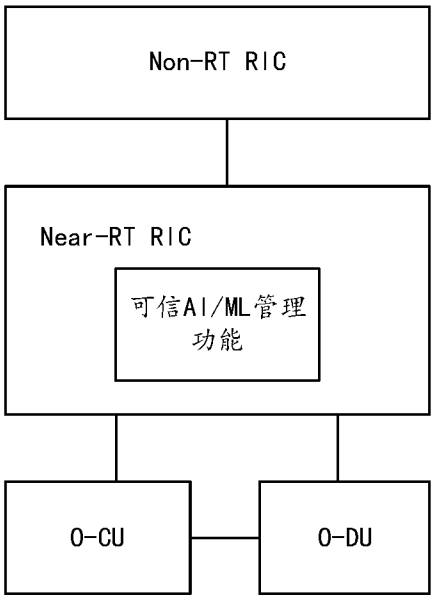

Non-RT RIC
Near-RT RIC
可信AI/ML管理
功能
O-CU
O-DU

图 2G

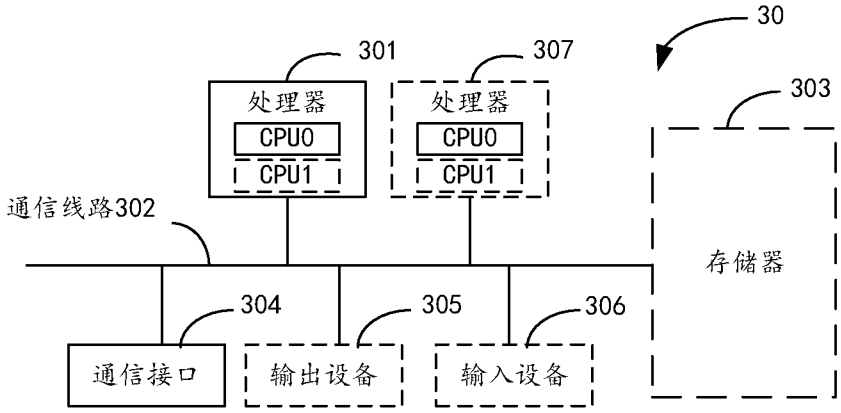

30
301
307
303
处理器
CPU0
CPU1
处理器
CPU0
CPU1
通信线路302
存储器
304
305
306
通信接口
输出设备
输入设备

图 3

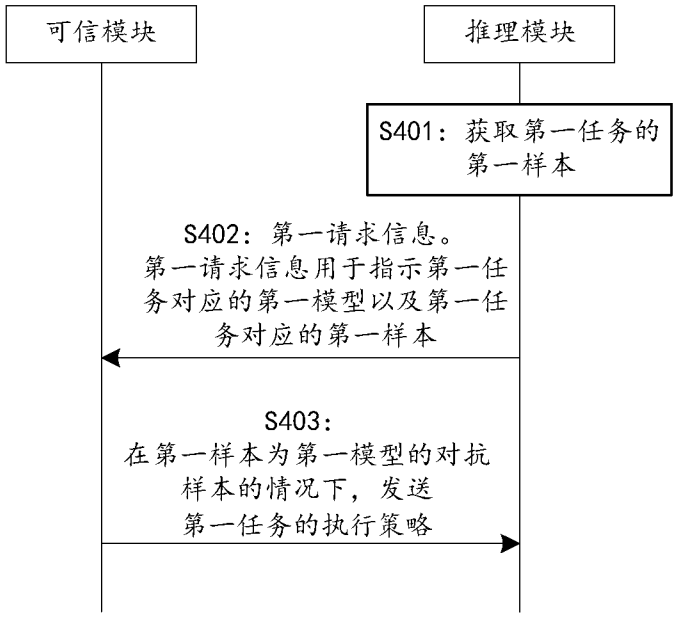

可信模块
推理模块
S401：获取第一任务的第一样本
S402：第一请求信息。第一请求信息用于指示第一任务对应的第一模型以及第一任务对应的第一样本
S403：在第一样本为第一模型的对抗样本的情况下，发送第一任务的执行策略

图 4

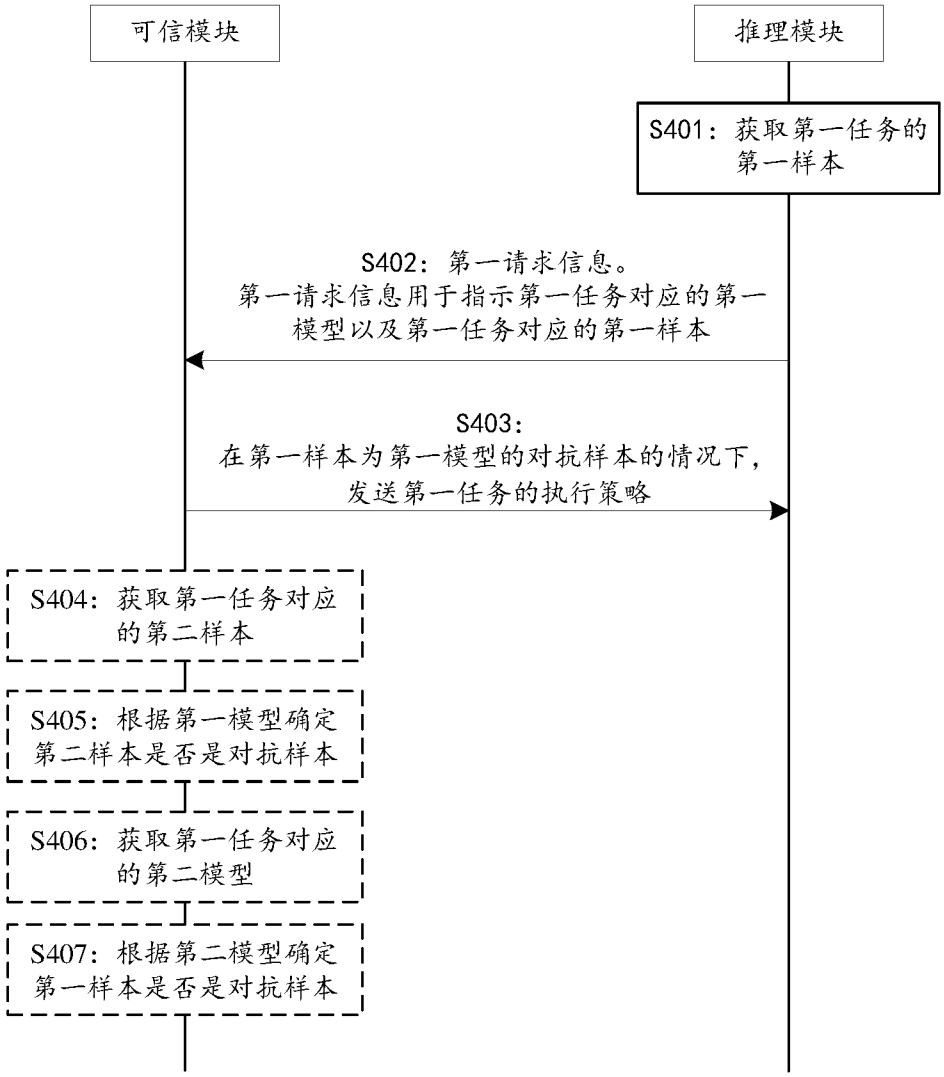

可信模块
推理模块
S401：获取第一任务的第一样本
S402：第一请求信息。
第一请求信息用于指示第一任务对应的第一模型以及第一任务对应的第一样本
S403：
在第一样本为第一模型的对抗样本的情况下，发送第一任务的执行策略
S404：获取第一任务对应的第二样本
S405：根据第一模型确定第二样本是否是对抗样本
S406：获取第一任务对应的第二模型
S407：根据第二模型确定第一样本是否是对抗样本

图 5

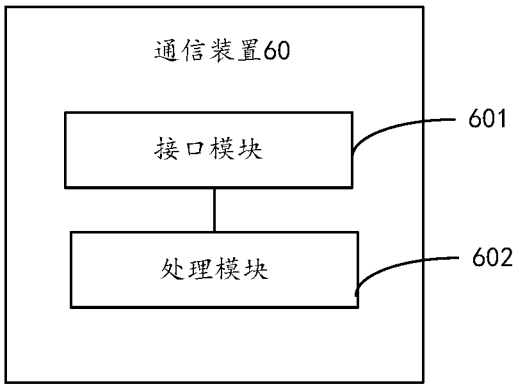

通信装置60
接口模块
601
处理模块
602

图 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2024/132767

A. CLASSIFICATION OF SUBJECT MATTER

G06F 9/48(2006.01)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:G06F

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, EPTXT, WOTXT, USTXT, VEN, WPABS, CNKI, IEEE: 策略, 对抗, 干扰, 扰动, 过滤, 平滑, 去噪, 压缩, 检测, 确定, 识别, 样本, 鲁棒, strategy, disturb, attack, sample, filter, smooth, denoise, detect, recognize, robust

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 115223127 A (CHONGQING CHANGAN AUTOMOBILE CO., LTD.) 21 October 2022 (2022-10-21) description, paragraphs 34-79	1-25
X	CN 115600107 A (ZHEJIANG DAHUA TECHNOLOGY CO., LTD.) 13 January 2023 (2023-01-13) description, paragraphs 55-110, and claims 1-25	1-25
A	CN 116304923 A (WUHAN UNIVERSITY) 23 June 2023 (2023-06-23) entire document	1-25
A	CN 110741388 A (DONGGUAN UNIVERSITY OF TECHNOLOGY) 31 January 2020 (2020-01-31) entire document	1-25
A	WO 2021143478 A1 (SHANGHAI FENGBAO INFORMATION TECHNOLOGY CO., LTD.) 22 July 2021 (2021-07-22) entire document	1-25

☐ 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	Date of mailing of the international search report
17 January 2025	**22 January 2025**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2024/132767

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 115223127 A	21 October 2022	None	
CN 115600107 A	13 January 2023	None	
CN 116304923 A	23 June 2023	None	
CN 110741388 A	31 January 2020	US 2021049505 A1	18 February 2021
		US 10936973 B1	02 March 2021
		WO 2021026805 A1	18 February 2021
WO 2021143478 A1	22 July 2021	None	

国际检索报告

国际申请号
PCT/CN2024/132767

A. 主题的分类

G06F 9/48(2006.01)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:G06F

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,EPTXT,WOTXT,USTXT,VEN,WPABS,CNKI,IEEE:策略,对抗,干扰,扰动,过滤,平滑,去噪,压缩,检测,确定,识别,样本,鲁棒,strategy,disturb,attack,sample,filter,smooth,denoise,detect,recognize,robust

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	CN 115223127 A (重庆长安汽车股份有限公司) 2022年10月21日 (2022 - 10 - 21) 说明书第34-79段	1-25
X	CN 115600107 A (浙江大华技术股份有限公司) 2023年1月13日 (2023 - 01 - 13) 说明书第55-110段 权1-25	1-25
A	CN 116304923 A (武汉大学) 2023年6月23日 (2023 - 06 - 23) 全文	1-25
A	CN 110741388 A (东莞理工学院) 2020年1月31日 (2020 - 01 - 31) 全文	1-25
A	WO 2021143478 A1 (SHANGHAI FENGBAO INFORMATION TECHNOLOGY CO., LTD.) 2021年7月22日 (2021 - 07 - 22) 全文	1-25

☐ 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	国际检索报告邮寄日期
2025年1月17日	2025年1月22日

ISA/CN的名称和邮寄地址	受权官员
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	李娜 电话号码 (+86) 010-53961403

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/132767

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	115223127	A	2022年10月21日	无			
CN	115600107	A	2023年1月13日	无			
CN	116304923	A	2023年6月23日	无			
CN	110741388	A	2020年1月31日	US	2021049505	A1	2021年2月18日
				US	10936973	B1	2021年3月2日
				WO	2021026805	A1	2021年2月18日
WO	2021143478	A1	2021年7月22日	无			