

(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
2024 年 5 月 16 日 (16.05.2024)



(10) 国际公布号
WO 2024/099111 A1

(51) 国际专利分类号:
H04W 4/44 (2018.01)

南 (YUAN, Yunnan); 中国广东省东莞市长安镇
维沃路1号, Guangdong 523863 (CN)。

(21) 国际申请号: PCT/CN2023/127307

(22) 国际申请日: 2023 年 10 月 27 日 (27.10.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202211387610.2 2022年11月7日 (07.11.2022) CN

(71) 申请人: 维沃移动通信有限公司 (VIVO MOBILE
COMMUNICATION CO., LTD.) [CN/CN]; 中国广东
省东莞市长安镇维沃路1号, Guangdong
523863 (CN)。

(72) 发明人: 周通 (ZHOU, Tong); 中国广东省东莞市长
安镇维沃路1号, Guangdong 523863 (CN)。 袁雁

(74) 代理人: 北京远志博慧知识产权代理事务
所 (特殊普通合伙) (BOHUI INTELLECTUAL
PROPERTY); 中国北京市海淀区交大东路31号东
区10号楼等17幢31幢108, Beijing 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,

(54) Title: AI SERVICE PROCESSING METHOD AND DEVICE

(54) 发明名称: AI服务的处理方法及设备

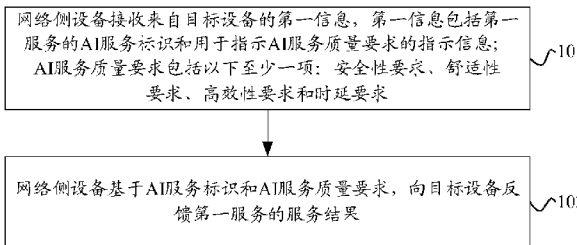


图 2

- 101 A network side device receives first information from a target device, the first information comprising an AI service identifier of a first service and indication information used for indicating an AI service quality requirement, the AI service quality requirement comprising at least one of the following: a safety requirement, a comfort requirement, a high efficiency requirement, and a delay requirement
- 102 The network side device feeds back the service result of the first service to the target device on the basis of the AI service identifier and the AI service quality requirement

(57) Abstract: The present application relates to the technical field of communications, and discloses an AI service processing method and a device. The AI service processing method in embodiments of the present application comprises: a network side device receives first information from a target device, the first information comprising an AI service identifier of a first service and indication information used for indicating an AI service quality requirement, and the AI service quality requirement comprising at least one of the following types of requirements: a safety requirement, a comfort requirement, a high efficiency requirement, and a delay requirement, wherein the safety requirement is used for representing the anti-collision capability of the target device, the comfort requirement is used for representing the riding comfort from the target device, the high efficiency requirement is used for representing the task completion efficiency of the target device, and the delay requirement is used for representing the difference between the time when the target device initiates an AI service request and the time when the target device obtains a service result; and the network side device feeds back the

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))。

service result of the first service to the target device on the basis of the AI service identifier and the AI service quality requirement.

(57) 摘要: 本申请公开了一种AI服务的处理方法及设备, 属于通信技术领域, 本申请实施例的AI服务的处理方法包括: 网络侧设备接收来自目标设备的第一信息, 所述第一信息包括第一服务的AI服务标识和用于指示AI服务质量要求的指示信息; AI服务质量要求包括以下至少一种类型的要求: 安全性要求、舒适性要求、高效性要求和时延要求; 安全性要求用于表征目标设备的防碰撞能力; 所述舒适性要求用于表征目标设备的乘坐舒适度; 高效性要求用于表征目标设备完成任务的效率; 时延要求用于表征目标设备发起AI服务请求与获取到服务结果之间的时间差; 网络侧设备基于AI服务标识和AI服务质量要求, 向目标设备反馈第一服务的服务结果。

AI服务的处理方法及设备

相关申请的交叉引用

本申请主张在 2022 年 11 月 07 日提交的申请号为 202211387610.2 的中国专利的
5 优先权，其全部内容通过引用包含于此。

技术领域

本申请属于通信技术领域，具体涉及一种 AI 服务的处理方法及设备。

背景技术

在移动通信系统中，现在有越来越多的用例结合人工智能（Artificial Intelligence，
10 AI）。例如在物理层有基于 AI 的 CSI（channel state information）反馈压缩，基于 AI 的波束管理，基于 AI 的定位。3GPP RAN3 讨论了基于 AI 的节能，基于 AI 的负载均衡等。在未来会有更多的结合 AI 的用例在移动通信系统中出现。

未来的移动通信系统不仅可以向内为通信系统提供 AI 服务。也可以向外部应用提供 AI 服务。例如，为无人驾驶、无人机、机器人等通信网络外部应用提供 AI 服务。
15 目前移动通信系统的无线网络只是作为一个通信管道，即只用于传输信息，而并不能为外部应用提供 AI 服务，因此如何更好地在移动通信系统中支持 AI 服务，是需要解决的问题。

发明内容

本申请实施例提供一种 AI 服务的处理方法及设备，能够解决如何更好地在移动通信系统中支持 AI 服务的问题。
20

第一方面，提供了一种 AI 服务的处理方法，包括：

网络侧设备接收来自目标设备的第一信息，所述第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；所述 AI 服务质量要求包括以下至少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；所述安全性要求
25 用于表征所述目标设备的防碰撞能力；所述舒适性要求用于表征所述目标设备的乘坐舒适度；所述高效性要求用于表征所述目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；

所述网络侧设备基于所述 AI 服务标识和 AI 服务质量要求，向目标设备反馈第一服务的服务结果。

第二方面，提供了一种 AI 服务的处理方法，包括：

目标设备向网络侧设备发送第一信息，所述第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；所述 AI 服务质量要求包括以下至少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；所述安全性要求用于
30 表征所述目标设备的防碰撞能力；所述舒适性要求用于表征所述目标设备的乘坐舒适度；所述高效性要求用于表征所述目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；

所述目标设备接收所述网络侧设备发送的第一服务的服务结果。

第三方面，提供了一种 AI 服务的处理装置，包括：

接收模块，用于接收来自目标设备的第一信息，所述第一信息包括第一服务的 AI
40 服务标识和用于指示 AI 服务质量要求的指示信息；所述 AI 服务质量要求包括以下至

少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；所述安全性要求用于表征所述目标设备的防碰撞能力；所述舒适性要求用于表征所述目标设备的乘坐舒适度；所述高效性要求用于表征所述目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；

5 处理模块，用于基于所述 AI 服务标识和 AI 服务质量要求，向目标设备反馈第一服务的服务结果。

第四方面，提供了一种 AI 服务的处理装置，包括：

发送模块，用于向网络侧设备发送第一信息，所述第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；所述 AI 服务质量要求包括以下至少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；所述安全性要求用于表征所述目标设备的防碰撞能力；所述舒适性要求用于表征所述目标设备的乘坐舒适度；所述高效性要求用于表征所述目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；

接收模块，用于接收所述网络侧设备发送的服务结果。

15 第五方面，提供了一种网络侧设备，该网络侧设备包括处理器和存储器，所述存储器存储可在所述处理器上运行的程序或指令，所述程序或指令被所述处理器执行时实现如第一方面所述的方法的步骤。

第六方面，提供了一种网络侧设备，包括处理器及通信接口，其中，所述通信接口用于接收来自目标设备的第一信息，所述第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；所述 AI 服务质量要求包括以下至少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；所述安全性要求用于表征所述目标设备的防碰撞能力；所述舒适性要求用于表征所述目标设备的乘坐舒适度；所述高效性要求用于表征所述目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；所述处理器用于基于所述 AI 服务标识和 AI 服务质量要求，向目标设备反馈第一服务的服务结果。

第七方面，提供了一种目标设备，该目标设备包括处理器和存储器，所述存储器存储可在所述处理器上运行的程序或指令，所述程序或指令被所述处理器执行时实现如第二方面所述的方法的步骤。

第八方面，提供了一种目标设备，包括处理器及通信接口，其中，所述通信接口用于向网络侧设备发送第一信息，所述第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；所述 AI 服务质量要求包括以下至少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；所述安全性要求用于表征所述目标设备的防碰撞能力；所述舒适性要求用于表征所述目标设备的乘坐舒适度；所述高效性要求用于表征所述目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；所述通信接口还用于接收所述网络侧设备发送的第一服务的服务结果。

第九方面，提供了一种通信系统，包括：目标设备及网络侧设备，所述目标设备可用于执行如第二方面所述的 AI 服务的处理方法的步骤，所述网络侧设备可用于执行如第一方面所述的 AI 服务的处理方法的步骤。

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

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

第十二方面，提供了一种计算机程序/程序产品，所述计算机程序/程序产品被存储在存储介质中，所述计算机程序/程序产品被至少一个处理器执行以实现如第一方面或第二方面所述的 AI 服务的处理方法的步骤。

在本申请实施例中，网络侧设备接收来自目标设备的第一信息，第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；AI 服务质量要求包括以下至少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；安全性要求用于表征所述目标设备的防碰撞能力；舒适性要求用于表征目标设备的乘坐舒适度；高效性要求用于表征目标设备完成任务的效率；时延要求用于表征目标设备发起 AI 服务请求与获取到服务结果之间的时间差；进一步地，网络侧设备基于 AI 服务标识和 AI 服务质量要求，获取第一服务的服务结果，并向目标设备反馈服务结果，即网络侧设备可以基于 AI 服务质量要求为目标设备提供服务，可以提升 AI 服务质量，从而实现了更好地在移动通信系统中支持 AI 服务。

附图说明

图 1 是本申请实施例可应用的无线通信系统的结构图；

图 2 是本申请实施例提供的 AI 服务的处理方法的流程示意图之一；

图 3 是本申请实施例提供的 AI 服务的处理方法的交互流程示意图之一；

图 4 是本申请实施例提供的 AI 服务的处理方法的流程示意图之二；

图 5 是本申请实施例提供的 AI 服务的处理装置的结构示意图之一；

图 6 是本申请实施例提供的 AI 服务的处理装置的结构示意图之二；

图 7 是本申请实施例提供的通信设备的结构示意图；

图 8 是本申请实施例提供的终端的结构示意图；

图 9 是本申请实施例的网络侧设备的结构示意图。

具体实施方式

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

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

值得指出的是，本申请实施例所描述的技术不限于长期演进型（Long Term Evolution, LTE）/LTE 的演进（LTE-Advanced, LTE-A）系统，还可用于其他无线通信系统，诸如码分多址（Code Division Multiple Access, CDMA）、时分多址（Time Division Multiple Access, TDMA）、频分多址（Frequency Division Multiple Access, FDMA）、正交频分多址（Orthogonal Frequency Division Multiple Access, OFDMA）、单载波频分多址（Single-carrier Frequency Division Multiple Access, SC-FDMA）和其他系统。本申请实施例中的术语“系统”和“网络”常被可互换地使用，所描述的技术

术既可用于以上提及的系统 and 无线电技术，也可用于其他系统和无线电技术。以下描述出于示例目的描述了新空口（New Radio, NR）系统，并且在以下大部分描述中使用 NR 术语，但是这些技术也可应用于 NR 系统应用以外的应用，如第 6 代（6th Generation, 6G）通信系统。

图 1 示出本申请实施例可应用的一种无线通信系统的框图。无线通信系统包括终端 11 和网络侧设备 12。其中，终端 11 可以是手机、平板电脑（Tablet Personal Computer）、膝上型电脑（Laptop Computer）或称为笔记本电脑、个人数字助理（Personal Digital Assistant, PDA）、掌上电脑、上网本、超级移动个人计算机（ultra-mobile personal computer, UMPC）、移动上网装置（Mobile Internet Device, MID）、增强现实（augmented reality, AR）/虚拟现实（virtual reality, VR）设备、机器人、可穿戴式设备（Wearable Device）、车载设备（VUE）、行人终端（PUE）、智能家居（具有无线通信功能的家居设备，如冰箱、电视、洗衣机或者家具等）、游戏机、个人计算机（personal computer, PC）、柜员机或者自助机等终端侧设备，可穿戴式设备包括：智能手表、智能手环、智能耳机、智能眼镜、智能首饰（智能手镯、智能手链、智能戒指、智能项链、智能脚镯、智能脚链等）、智能腕带、智能服装等。需要说明的是，在本申请实施例并不限定终端 11 的具体类型。网络侧设备 12 可以包括接入网设备或核心网设备，其中，接入网设备 12 也可以称为无线接入网设备、无线接入网（Radio Access Network, RAN）、无线接入网功能或无线接入网单元。接入网设备 12 可以包括基站、WLAN 接入点或 WiFi 节点等，基站可被称为节点 B、演进节点 B（eNB）、接入点、基收发机站（Base Transceiver Station, BTS）、无线电基站、无线电收发机、基本服务集（Basic Service Set, BSS）、扩展服务集（Extended Service Set, ESS）、家用 B 节点、家用演进型 B 节点、发送接收点（Transmitting Receiving Point, TRP）或所述领域中其他某个合适的术语，只要达到相同的技术效果，所述基站不限于特定技术词汇，需要说明的是，在本申请实施例中仅以 NR 系统中的基站为例进行介绍，并不限定基站的具体类型。核心网设备可以包含但不限于如下至少一项：核心网节点、核心网功能、移动管理实体（Mobility Management Entity, MME）、接入移动管理功能（Access and Mobility Management Function, AMF）、会话管理功能（Session Management Function, SMF）、用户平面功能（User Plane Function, UPF）、策略控制功能（Policy Control Function, PCF）、策略与计费规则功能单元（Policy and Charging Rules Function, PCRF）、边缘应用服务发现功能（Edge Application Server Discovery Function, EASDF）、统一数据管理（Unified Data Management, UDM）、统一数据仓储（Unified Data Repository, UDR）、归属用户服务器（Home Subscriber Server, HSS）、集中式网络配置（Centralized network configuration, CNC）、网络存储功能（Network Repository Function, NRF）、网络开放功能（Network Exposure Function, NEF）、本地 NEF（Local NEF, 或 L-NEF）、绑定支持功能（Binding Support Function, BSF）、应用功能（Application Function, AF）等。需要说明的是，在本申请实施例中仅以 NR 系统中的核心网设备为例进行介绍，并不限定核心网设备的具体类型。

未来的移动通信系统不仅可以向内为通信系统提供 AI 服务。也可以向外部应用提供 AI 服务。例如，为无人驾驶、无人机、机器人等通信网络外部应用提供 AI 服务。这些千行百业的外部应用有不同的服务质量等级。如果还基于相关通信系统中的服务质量指标作为外部服务的服务指标。则无法匹配外部应用的服务质量要求。因此，本申请实施例中提供了一种 AI 服务的处理方法，在发送 AI 服务请求时携带用于指示 AI

服务质量要求的指示信息,使得移动通信系统可以接收 AI 服务质量要求,使得移动通信系统基于 AI 服务质量要求可以为外部应用提供质量更高的服务。

下面结合附图,通过一些实施例及其应用场景对本申请实施例提供的 AI 服务的处理方法进行详细地说明。

5 图 2 是本申请实施例提供的 AI 服务的处理方法的流程示意图之一。如图 2 所示,本实施例提供的方法,包括:

步骤 101、网络侧设备接收来自目标设备的第一信息,第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息; AI 服务质量要求包括以下至少一种类型的要求: 安全性要求、舒适性要求、高效性要求和时延要求; 安全性要求用于表征所述目标设备的防碰撞能力; 舒适性要求用于表征目标设备的乘坐舒适度; 高效性要求用于表征目标设备完成任务的效率; 时延要求用于表征目标设备发起 AI 服务请求与获取到服务结果之间的时间差;

具体地,网络侧设备例如基站功能或核心网功能,目标设备例如终端,车辆。

15 目标设备向网络侧设备发送第一信息,例如第一信息可以是 AI 服务请求,网络侧设备接收该第一信息,第一信息包括: AI 服务标识和用于指示 AI 服务质量要求的指示信息。

例如,目标设备请求的服务为环境感知,例如对车道线、静态物体和动态物体等的感知,或者目标设备请求的服务为路径规划,例如提供一条轨迹,包含多个轨迹点标识,每个轨迹点标识对应一个距离和速度; 目标设备例如为车载终端,安全性要求可以通过车辆与障碍物的距离计算的代价表示,舒适性要求可以通过车辆行驶的路径朝向偏差、加速度等参数计算的代价表示,高效性要求用于表征目标设备完成任务的效率,例如通过路径行驶时长、与理想无障碍路径偏差、是否超过道路边界等参数计算的代价表示,时延要求用于表征目标设备发起 AI 服务请求与获取到网络侧设备反馈的服务结果之间的时间差,例如时延要求为小于 100ms。

25 步骤 102、网络侧设备基于 AI 服务标识和 AI 服务质量要求,向目标设备反馈第一服务的的服务结果。

具体地,网络侧设备根据 AI 服务标识和 AI 服务质量要求,提供相应的服务,即获取该 AI 服务标识对应的第一服务的的服务结果,例如基于 AI 模型获取服务结果,例如可以根据 AI 服务质量要求对 AI 模型的输出进行迭代,获取服务质量较好的服务结果,并把服务结果反馈给目标设备。

35 本实施例的方法,网络侧设备接收来自目标设备的第一信息,第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息; AI 服务质量要求包括以下至少一种类型的要求: 安全性要求、舒适性要求、高效性要求和时延要求; 安全性要求用于表征所述目标设备的防碰撞能力; 舒适性要求用于表征目标设备的乘坐舒适度; 高效性要求用于表征目标设备完成任务的效率; 时延要求用于表征目标设备发起 AI 服务请求与获取到服务结果之间的时间差; 进一步地,网络侧设备基于 AI 服务标识和 AI 服务质量要求,获取第一服务的的服务结果,并向目标设备反馈服务结果,即网络侧设备可以基于 AI 服务质量要求为目标设备提供服务,可以提升 AI 服务质量,从而实现了更好地在移动通信系统中支持 AI 服务。

40 可选地,网络侧设备获取用于确定质量评估方式的信息; 用于确定质量评估方式的信息包括: 质量评估方式对应的可执行文件和运行信息; 质量评估方式用于对服务结果的质量进行评估。

可选地，运行信息包括以下至少一项：

运行环境配置信息、文件名称、输入格式、输出格式、指示预处理方式的信息、指示后处理方式的信息。

可选地，网络侧设备获取用于确定质量评估方式的信息，包括以下至少一项：

5 网络侧设备接收来自目标设备的用于确定质量评估方式的信息；

网络侧设备接收来自目标设备的第三方链接，并基于第三方链接获取用于确定质量评估方式的信息；

网络侧设备接收来自目标设备的应用程序编程接口（Application Programming Interface, API），并基于 API 获取用于确定质量评估方式的信息。

10 具体地，目标设备可以将用于确定质量评估方式的信息直接发送给网络侧设备，或通过第三方链接发送给网络侧设备，或通过 API 发送给网络侧设备，网络侧设备直接从接收到的信息中获取用于确定质量评估方式的信息，或，基于第三方链接或 API 获取用于确定质量评估方式的信息。

15 可选地，网络侧设备可以基于用于确定质量评估方式的信息，对服务结果进行评价，获取评价结果。

例如，还可以基于评价结果以及服务质量要求对提供的服务进行优化，例如优化 AI 模型。

20 可选地，在 AI 服务质量要求包括安全性要求的情况下，第一信息包括安全性要求指标门限和用于评估安全性质量的第一参数信息；其中，第一参数信息包括以下至少一项：

第一指示信息、第二信息、第一加权信息；

其中，第一指示信息用于指示安全性质量评估包括的至少一个第一目标参数项；第一加权信息包括各个所述第一目标参数项的加权值；第二信息为各个所述第一目标参数项与安全代价的函数映射信息，或者第一距离与安全等级的对应关系；

25 所述第一目标参数项包括以下至少一项：

第一距离、第一保护间隔门限、第二保护间隔门限；

其中，第一距离是目标设备与障碍物的距离；第一保护间隔门限用于表示最高安全等级的保护间隔下限，第二保护间隔门限用于表示最低安全等级的保护间隔下限。

30 具体地，函数映射信息例如用于指示与第一目标参数项对应的安全代价函数的标识，或安全代价函数的内容。

当第一距离大于保护间隔第一门限时，安全等级属于最高等级。当第一距离小于保护间隔第二门限时，会发生安全事故，不满足服务质量要求。

例如，评估安全性质量，即计算安全代价可以通过第一距离、第一保护间隔门限、第二保护间隔门限、第一距离与安全等级的对应关系等实现。

35 可选地，在第一距离位于第一保护间隔门限和第二保护间隔门限之间时，第一距离越大，安全等级越高。

例如，安全等级的数值越大，表示安全性越高；或，安全等级的数值越小，表示安全性越高。

40 在安全等级的数值越小，表示安全性越高的情况下，当第一距离位于第一门限和第二门限之间时，安全等级是第一距离的单调递减函数。

可选地，网络侧设备根据以下至少一项信息，确定第一距离；

至少一项信息包括：目标设备的尺寸、航向角、目标设备的位置和障碍物边界。

具体地，目标设备例如为车辆，航向角为该车辆的航向角。

示例性地，以无人车为例，第一距离为无人车边界到障碍物边界之间的距离。例如，无人车边界以覆盖无人车的三维度长方体来表示。同理，障碍物的边界也是由覆盖障碍物的三维长方体来表示。

5 第一保护间隔门限为 10 米，表示当第一距离大于或等于 10 米时，安全代价为 0；
 第二保护间隔门限为 0.1 米，表示当第一距离小于 0.1 米时，规划的轨迹无效。安全代价为负无穷大。

安全代价的函数标识为 4，当第一距离大于或等于 0.1 米，且小于 10 米时，选择函数 $\exp(-d/10)$ ，d 为第一距离，来计算安全代价，如下表 1 所示：

10

表 1

函数标识	第一距离对应的安全代价的函数
1	$\exp(-d/2)$
2	$\exp(-d/4)$
3	$\exp(-d/8)$
4	$\exp(-d/10)$

可选地，在 AI 服务质量要求包括舒适性要求的情况下，第一信息包括舒适性要求指标门限和用于评估舒适性质量的第二参数信息；其中，第二参数信息包括以下至少一项：

第二指示信息、第三信息、第二加权信息；

15 其中，第二指示信息用于指示舒适性质量评估包括的至少一个第二目标参数项，第三信息为各个所述第二目标参数项对应的舒适性代价函数的信息；第二加权信息包括各个所述第二目标参数项的加权值；所述第二目标参数项包括以下至少一项：

路径的朝向偏差、路径的航向角、路径的曲率、路径的曲率导数、路径的曲率变化率、加速度、加加速度。

20 具体地，第二指示信息用于指示用于舒适性质量评估的包括的目标参数项，例如，舒适性质量评估时考虑以下参数：路径的朝向偏差、航向角、曲率。

示例性地，第一信息中包括朝向偏差代价函数的信息（例如朝向偏差代价函数标识）、路径的曲率代价函数的信息（函数标识）和曲率变化率代价函数的信息（函数标识）。而这三者的代价函数是提前约定好的。三者的加权值也是提前约定的。如果
25 三者的加权值与约定值不同，也可以携带在第一信息中，即通过第一信息指示。

例如，舒适性评估考虑参数：路径的朝向偏差、路径的曲率、曲率变化率；

假设总规划时长是 15 秒，时间间隔是 1 秒，则这个路径包括规划的 15 个轨迹点。

30 路径的朝向偏差为路径的一阶导数，朝向偏差代价的计算方法为，在每个时间间隔内，对路径的一阶导数求平方；再把 15 个时间间隔的值累加起来。计算方法可以是预定义的，也可以通过包含在第二信息的朝向偏差代价函数信息中。

路径的曲率为路径的二阶导数，路径的曲率代价的计算方法为，在每个时间间隔内，对路径的二阶导数求平方；再把 15 个时间间隔的值累加起来。

曲率变化率为路径的三阶导数，曲率变化率代价的计算方法为，在每个时间间隔内，对路径的三阶导数求平方；再把 15 个时间间隔的值累加起来。

35 朝向偏差代价、路径的曲率代价和曲率变化率代价的加权值分别为 w1, w2 和 w3。则舒适度总代价为：

$w1 \times \text{朝向偏差代价} + w2 \times \text{路径的曲率代价} + w3 \times \text{曲率变化率代价}$ 。

可选地，在 AI 服务质量要求包括所述高效性要求的情况下，第一信息求包括所述高效性要求指标门限和用于评估高效性质量的第三参数信息；其中，第三参数信息包括以下至少一项：

5 第三指示信息、第四信息、第三加权值；

其中，第三指示信息用于指示高效性质量评估包括的至少一个第三目标参数项，第四信息为各个所述第三目标参数项对应的高效性代价函数的信息；所述第三加权信息包括各个所述第三目标参数项的加权值；所述第三目标参数项包括以下至少一项：

10 与车道中心线的偏差、与理想无障碍路径的偏差、超过道路边界的惩罚值、路径行驶时长、路径行驶距离。

具体地，第三指示信息用于指示用于高效性质量评估的包括的目标参数项，例如，高效性质量评估时考虑以下参数：与车道中心线的偏差、路径行驶时长、路径行驶距离、超过道路边界的惩罚值。示例性地，第一信息中包括以下几项：

15 与车道中心线的偏差代价函数信息；（表示与理想无障碍路径偏差的惩罚函数是提前约定好的）

超过道路边界的惩罚值为 100；（表示与预先约定的值不同，或者预先没有约定该值）

路径行驶时长代价函数信息；（表示已经提前约定了计算方法）

20 与车道中心线的偏差代价、超过道路边界的惩罚值和路径行驶时长代价的加权值分别为 $w1$ ， $w2$ 和 $w3$ 。

假设，总规划时长是 15 秒，时间间隔是 1 秒，则这个路径包括规划的 15 个轨迹点。

25 与理想无障碍路径偏差代价的计算方法为，在每个时间间隔内，对规划路径与理想无障碍路径的差值求平方；再把 15 个时间间隔的值累加起来。计算方法可以是预定义的，也可以通过包含在第三信息的朝向偏差代价函数信息中。

路径行驶时长代价的计算方法为 $1/(1+\text{第二距离})$ ，其中第二距离为规划路径的总行驶距离。计算方法可以是预定义的，也可以通过包含在第三信息的朝向偏差代价函数信息中。

则高效性要求的总代价为：

30 $w1 \times \text{与理想无障碍路径偏差代价} + w2 \times 100 + w3 \times 1/(1+\text{第二距离})$ 。

可选地，AI 服务质量要求还包括：

用于指示服务质量要求类型的指示信息，以及每个类型的服务质量要求的合并方式，所述合并方式包括以下至少一项：

每个类型的服务质量要求包括的参数项合并的加权函数信息；

35 每个类型的服务质量要求包括的参数项的加权值。

具体地，加权函数信息例如包括：加权函数的标识或加权函数的内容，例如安全性要求、舒适性要求、高效性要求和时延要求分别对应的加权函数，每一种类型中多个参数对应的加权值，可以获取到每个类型的服务质量要求。

可选地，综合服务质量要求为由第一信息确定的，包括以下至少一种方式：

40 基于第一信息直接得到的；

基于第一信息得到 AI 服务质量要求，再基于 AI 服务质量要求得到的。

具体地，综合服务质量要求可以由第一信息直接得到，例如第一信息中直接包

括综合服务质量要求的内容，或包括计算综合服务质量要求的内容；或，先由第一信息得到 AI 服务质量要求，再基于 AI 服务质量要求得到的，例如第一信息包括多项服务质量要求，由该多项服务质量要求得到综合服务质量要求。

可选地，在综合服务质量要求为由第一信息确定的情况下，第一信息还包括以下至少一项：

- 综合服务质量等级；
- 综合服务质量门限；
- 各服务质量的权重；
- 各服务质量的加权方法。

可选地，综合服务质量等级与各服务质量等级对应关系，例如表 2 所示：

表 2

综合服务质量等级	舒适性	安全性	高效性	时延
1	1	1	1	1
2	1	1	2	1
3	2	1	2	1
4	3	1	2	1

可选地，综合服务质量等级与各服务质量数值（服务质量数值通过代价或时延表示）的对应关系，例如表 3 所示：

表 3

综合服务质量等级	舒适性代价	安全性代价	高效性代价	时延
1	<20	<5	<10	<100ms
2	<20	<5	<50	<100ms
3	<50	<5	<50	<100ms
4	<100	<5	<50	<100ms

示例性地，综合服务质量值可以由各服务质量数值映射得到。例如，涉及的质量要求指标包括安全性，高效性，舒适性和时延。

可选地，综合质量值计算方式如下：

综合服务质量值 = $a_1 \times \text{安全性服务质量数值} + a_2 \times \text{高效性服务质量数值} + a_3 \times \text{舒适性服务质量数值} + a_4 \times \text{时延}$ 。其中， a_1 ， a_2 ， a_3 和 a_4 分别为四者加权值。

可选地，综合服务质量等级与舒适性代价的对应关系，例如表 4 所示：

表 4

综合服务质量等级	舒适性代价
1	小于等于 20
2	大于 20 且小于等于 50
3	大于 50 且小于等于 100

可选地，综合服务质量等级与安全性代价的对应关系，例如表 5 所示：

表 5

综合服务质量等级	防碰撞代价
1	小于等于 5
2	大于 5 且小于等于 10
3	大于 10 且小于等于 20

可选地，综合服务质量等级与高效性代价的对应关系，例如表 6 所示：

表 6

综合服务质量等级	高效性代价
1	小于等于 10
2	大于 10 且小于等于 50
3	大于 50 且小于等于 100

可选地，综合服务质量等级与时延的对应关系，例如表 7 所示：

表 7

综合服务质量等级	时延
1	小于等于 100ms
2	大于 100ms 且小于等于 150
3	大于 150ms 且小于等于 200

- 5
- 10
- 15
- 20
- 25
- 30
- 可选地，AI 服务质量要求通过以下至少一项表示：
- 质量等级、质量要求指标门限；质量要求指标门限包括以下至少一项：安全性要求指标门限、舒适性要求指标门限、高效性要求指标门限和时延要求指标门限。
- 可选地，AI 服务质量要求还包括以下至少一种类型的要求：
- 速度保持代价；
- 路径单调性；
- 目标设备速度上限；
- 路径规划的总时长；
- 路径规划的时间间隔；
- 其中，速度保持代价为基于服务结果中包括的行驶速度与目标行驶速度的差值得到的，目标行驶速度为由第二信息确定的；第二信息包括以下至少一项：路径的曲率、交通规则；
- 路径单调性是指在一次规划的路径中目标设备与路径起点的距离是单调递增的，或目标设备与路径终点的距离是单调递减的；
- 目标设备速度上限为基于交通规则和/或目标设备动力学得到的。
- 具体地，速度保持代价可以通过网络侧设备计算得到的行驶速度，与目标行驶速度的差值得到的，偏离目标行驶速度会受到惩罚。目标行驶速度由路径的曲率、交通规则等确定，例如基于道路速度限制、曲率和其他交通规则等来确定。上述速度保持代价、路径单调性、目标设备速度上限等可用于网络侧设备为目标设备提供服务，例如上述信息作为网络侧设备的 AI 模型的输入数据，可以提高服务质量。
- 可选地，第一信息中还包括：用于计算速度保持代价的信息，用于计算速度保持代价的信息包括以下至少一项：
- 偏离速度惩罚函数指示信息；
- 偏离速度的惩罚倍数。
- 可选地，偏离速度惩罚函数指示信息，例如为偏离速度惩罚函数的具体内容，或函数标识。惩罚倍数可以为计算速度保持代价的权值。例如，偏离速度惩罚函数为 f ，则速度保持代价为 $f(\text{偏离速度})$ ，或者惩罚倍数为 α_p ，则速度保持代价为 $\alpha_p \times \text{偏离速度}$ ，其中， f 例如为服务结果中包括的行驶速度与目标行驶速度的差值的函数。
- 可选地，第一信息中还包括以下至少一项：

目标设备的尺寸;

目标设备的位置;

目标设备对周围物体的感知结果;感知结果包括以下至少一项的感知结果:类型、距离、速度、加速度、边界;

5 目标设备对周围物体的原始感知数据。

具体地,目标设备例如为车辆,通过第一信息中包括的目标设备的尺寸、目标设备的位置、原始感知数据、感知结果等,可以辅助网络侧设备进行 AI 模型的推理或计算,以提升网络侧设备提供的服务质量。

可选地,该方法还包括:

10 网络侧设备基于第一信息和质量评估方式的信息,或,基于第一信息确定第六信息,第六信息包括以下至少一项:质量要求指标、质量要求指标计算方法、质量要求指标输入数据和格式、质量要求指标门限。

具体地,质量要求指标例如包括以下至少一项:安全性、舒适性、高效性和时延指标,基于第六信息可以获取服务结果,为目标设备提供服务。或者,基于第六信息,15 还可以获取评价结果。

可选地,该方法还包括:

网络侧设备基于第六信息和服务结果,确定服务结果的第一评价结果;

网络侧设备向所述目标设备发送服务结果的第一评价结果。

具体地,网络侧设备获取到服务结果之后,可以发送给目标设备,可选地,还可20 以将服务结果的第一评价结果发送给目标设备,例如目标设备可以基于该第一评价结果对服务结果进行筛选,选择满足要求的服务结果,基于服务结果执行相应的任务。

可选地,目标设备还可以向网络侧设备发送服务结果的第二评价结果,网络侧设备接收目标设备发送的第二评价结果,第二评价结果为所述目标设备对服务结果进行评价的评价结果。

25 第二评价结果可以用于网络侧设备优化提供的服务,例如,网络侧设备可以基于目标设备反馈的第二评价结果对 AI 模型进行优化。

可选地,第一评价结果或第二评价结果包括以下至少一项:

AI 服务标识;

AI 服务标识对应的服务等级;

30 AI 服务标识对应的服务质量数值;

综合服务质量数值;

综合服务质量等级。

具体地,例如,AI 服务标识表示当前服务为环境感知,当前评价结果例如为该环境感知的服务结果的服务等级为 5 级,例如服务质量要求中服务等级为 4 级,数值越35 小,等级越高,说明提供的服务结果不满足服务质量要求,需要重新进行优化。其它参数类似。

可选地,步骤 102 可以通过如下方式实现:

网络侧设备基于 AI 服务标识和第七信息,确定提供服务的算法信息;第七信息包括以下至少一项:AI 服务质量要求、综合服务质量要求;

40 网络侧设备基于 AI 服务标识以及算法信息,向目标设备反馈服务结果。

具体地,网络侧设备基于 AI 服务标识和第七信息,确定提供服务的算法信息,基于算法信息执行服务算法(例如 AI 模型推理),获取服务结果,并向目标设备反馈服

务结果。

可选地，在提供的 AI 服务为训练服务的情况下，算法信息包括训练的奖励函数；在提供的 AI 服务为推理服务的情况下，算法信息包括推理模型。

可选地，该方法还包括：

5 网络侧设备基于 AI 服务标识、推理服务对应的服务结果和综合服务质量要求，确定推理服务对应的服务结果的第一评价结果。

例如推理服务对应的服务结果包括：目标设备预设时间范围内的速度、路径等，综合服务质量要求包括综合服务质量等级的要求，利用该速度和目标行驶速度计算速度保持代价，基于速度保持代价可以确定该服务结果的综合服务质量等级，例如速度保持代价越小，综合服务质量等级越高，假设速度保持代价对应的综合服务质量等级为3级，但是综合服务质量等级的要求为5级，此时假设安全等级的数值越大，等级越高，则说明不满足综合服务质量要求。

示例性地，目标设备为车辆，网络侧设备为基站，当前服务为轨迹规划服务。

服务质量要求已经预先定义好，无需额外指示。可以根据服务标识直接获得。

15 如图 3 所示, 该方法包括如下步骤:

步骤 1: 车辆向基站发送服务标识, 该服务标识用于指示路径规划服务。

步骤2: 基站根据服务标识, 获取第六信息。第六信息包括以下至少一项: 质量要求指标, 质量要求指标计算方法, 质量要求指标输入数据和格式, 质量要求指标门限。

20 例如，确定质量要求指标包括安全性、舒适性、高效性和时延。

质量要求指标计算方法，质量要求指标输入数据和格式已经预先定义好。

质量要求指标门限为综合服务质量门限，例如门限为 10。当门限小于 10 时，则满足服务质量要求。

25 步骤 3: 基于服务标识, 确定服务算法。执行服务算法 (例如 AI 模型推理), 获取服务结果, 再基于质量要求指标计算方法验证服务结果是否满足服务要求。或选择服务质量最高的一个服务结果。

步骤4: 反馈服务结果和评价结果。评价结果可以是一个综合服务质量值, 或者一个等级。

可选地,网络侧设备根据第一信息,确定综合服务质量要求,进而确定服务算法。

30 此时，质量要求指标，质量要求指标计算方法，质量要求指标输入数据和格式都是预先定义的。

例如，质量要求指标包括：舒适性、高效性、安全性和时延

质量要求指标计算方法，质量要求指标输入数据和格式都是预先定义的。如表 8 所示：

表 8

例子	第一信息包括	综合服务质量门限	综合服务质量要求中各指标加权方法
1	综合服务质量等级	预定义	预定义
2	综合服务质量门限	第一信息指示	预定义
3	综合服务质量门限,各服务质量的权重和加	第一信息指示	第一信息指示

	权方法		
--	-----	--	--

上述实施方式中，通过提供 AI 服务质量要求，使得网络侧设备可以获得更完整更丰富的信息，从而提升服务质量，目标设备通过享受网络侧设备提供的服务，降低了自身设备的复杂度，进一步，网络侧设备由于可以识别不同的 AI 服务质量要求，从而可以为目标设备提供更灵活的服务。

5 图 4 是本申请实施例提供的 AI 服务的处理方法的流程示意图之二。如图 3、图 4 所示，本实施例提供的方法，包括：

步骤 201、目标设备向网络侧设备发送第一信息，第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；AI 服务质量要求包括以下至少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；

10 其中，安全性要求用于表征目标设备的防碰撞能力；所述舒适性要求用于表征所述目标设备的乘坐舒适度；所述高效性要求用于表征所述目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；

步骤 202、目标设备接收网络侧设备发送的第一服务的服务结果。

15 可选地，所述指示信息通过以下至少一项表示所述 AI 服务质量要求：

质量等级、质量要求指标门限；其中，所述质量要求指标门限包括以下至少一项：安全性要求指标门限、舒适性要求指标门限、高效性要求指标门限和时延要求指标门限。

可选地，所述方法还包括：

20 所述目标设备向所述网络侧设备发送用于确定质量评估方式的信息；所述质量评估方式用于对服务结果的质量进行评估；所述用于确定质量评估方式的信息包括：质量评估方式对应的可执行文件和运行信息。

可选地，所述运行信息包括以下至少一项：

25 运行环境配置信息、文件名称、输入格式、输出格式、指示预处理方式的信息、指示后处理方式的的信息。

可选地，所述目标设备向所述网络侧设备发送质量评估方式的信息，包括以下至少一项：

所述目标设备向所述网络侧设备发送第三方链接，所述第三方链接用于获取所述用于确定质量评估方式的信息；

30 所述目标设备向所述网络侧设备发送应用程序编程接口 API，所述应用程序编程接口 API 用于获取所述用于确定质量评估方式的信息。

可选地，在所述 AI 服务质量要求包括所述安全性要求的情况下，所述第一信息包括所述安全性要求指标门限和用于评估安全性质量的第一参数信息；其中，第一参数信息包括以下至少一项：

35 第一指示信息、第二信息、第一加权信息；

其中，第一指示信息用于指示舒适性质量评估包括的至少一个第一目标参数项；所述第一加权信息包括各个所述第一目标参数项的加权值；

所述第一目标参数项包括以下至少一项：

第一距离、第一保护间隔门限、第二保护间隔门限；

40 其中，第一距离是所述目标设备与障碍物的距离；所述第一保护间隔门限用于表

示最高安全等级的保护间隔下限，所述第二保护间隔门限用于表示最低安全等级的保护间隔下限；

第二信息为各个所述第一目标参数项与安全代价的函数映射信息，或者第一距离与安全等级的对应关系。

5 可选地，在所述第一距离位于第一保护间隔门限和第二保护间隔门限之间时，安全等级越高，所述第一距离越大。

可选地，所述第一距离为基于以下至少一项信息确定的；

所述至少一项信息包括：目标设备的尺寸、航向角、目标设备的位置和障碍物边界。

10 可选地，在所述 AI 服务质量要求包括所述舒适性要求的情况下，所述第一信息包括所述舒适性要求指标门限和用于评估舒适性质量的第二参数信息；其中，所述第二参数信息包括以下至少一项：

第二指示信息、第三信息、第二加权信息；

其中，第二指示信息用于指示舒适性质量评估包括的至少一个第二目标参数项，

15 第三信息为各个所述第二目标参数项对应的舒适性代价函数的信息；所述第二加权信息包括各个所述第二目标参数项的加权值；所述第二目标参数项包括以下至少一项：

路径的朝向偏差、路径的航向角、路径的曲率、路径的曲率导数、路径的曲率变化率、加速度、加加速度。

20 可选地，在所述 AI 服务质量要求包括所述高效性要求的情况下，所述第一信息包括所述高效性要求指标门限和用于评估高效性质量的第三参数信息；其中，所述第三参数信息包括以下至少一项：

第三指示信息、第四信息、第三加权信息；

其中，第三指示信息用于指示高效性质量评估包括的至少一个第三目标参数项，

25 第四信息为各个所述第三目标参数项对应的高效性代价函数的信息；所述第三加权信息包括各个所述第三目标参数项的加权值；所述第三目标参数项包括以下至少一项：

与车道中心线的偏差、与理想无障碍路径的偏差、超过道路边界的惩罚值、路径行驶时长、路径行驶距离。

可选地，所述 AI 服务质量要求还包括以下至少一种类型的要求：

速度保持代价；

30 路径单调性；

目标设备速度上限；

路径规划的总时长；

路径规划的时间间隔；

35 其中，速度保持代价为基于服务结果中包括的行驶速度与目标行驶速度的差值得到的，目标行驶速度为由第五信息确定的；所述第五信息包括以下至少一项：路径的曲率、交通规则；

路径单调性是指在一次规划的路径中所述目标设备与路径起点的距离是单调递增的，或所述目标设备与路径终点的距离是单调递减的；

目标设备速度上限为基于交通规则和/或目标设备动力学得到的。

40 可选地，所述第一信息中还包括：用于计算速度保持代价的信息，所述用于计算速度保持代价的信息包括以下至少一项：

偏离速度惩罚函数信息；

偏离速度的惩罚倍数。

可选地，所述 AI 服务质量要求还包括用于指示服务质量要求类型的指示信息，以及每个类型的服务质量要求的合并方式，所述合并方式包括以下至少一项：

每个类型的服务质量要求包括的参数项合并的加权函数信息；

5 每个类型的服务质量要求包括的参数项的加权值。

可选地，第一信息中还包括以下至少一项：

目标设备的尺寸；

目标设备的位置；

目标设备对周围物体的感知结果；所述感知结果包括以下至少一项的感知结果：

10 类型、距离、速度、加速度、边界；

目标设备对周围物体的原始感知数据。

可选地，所述方法还包括：

所述目标设备接收所述网络侧设备发送的所述服务结果的第一评价结果；所述服务结果的第一评价结果为基于所述第六信息和所述服务结果确定的；所述第六信息包
15 括以下至少一项：质量要求指标、质量要求指标计算方法、质量要求指标输入数据和格式、质量要求指标门限。

可选地，所述方法还包括：

所述目标设备向所述网络侧设备发送第二评价结果，所述第二评价结果为所述目标设备对所述服务结果进行评价的评价结果。

20 可选地，第一评价结果或第二评价结果包括以下至少一项：

AI 服务标识；

AI 服务标识对应的服务等级；

AI 服务标识对应的服务质量数值；

综合服务质量数值；

25 综合服务质量等级。

可选地，综合服务质量要求为由所述第一信息确定的，包括以下至少一种方式：

基于所述第一信息直接得到的；

基于所述第一信息得到 AI 服务质量要求，再基于所述 AI 服务质量要求得到的。

30 可选地，在所述综合服务质量要求为由所述第一信息确定的情况下，所述第一信息还包括以下至少一项：

综合服务质量等级；

综合服务质量门限；

各服务质量的权重；

各服务质量的加权方法。

35 本实施例的方法，其具体实现过程与技术效果与网络侧方法实施例中类似，具体可以参见终端侧方法实施例中的详细介绍，此处不再赘述。

本申请实施例提供的 AI 服务的处理方法，执行主体可以为 AI 服务的处理装置。本申请实施例中以 AI 服务的处理装置执行 AI 服务的处理方法为例，说明本申请实施例提供的 AI 服务的处理装置。

40 图 5 是本申请提供的 AI 服务的处理装置的结构示意图之一。如图 5 所示，本实施例提供的 AI 服务的处理装置，包括：

接收模块 210，用于接收来自目标设备的第一信息，所述第一信息包括第一服务

的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；所述 AI 服务质量要求包括以下至少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；所述安全性要求用于表征所述目标设备的防碰撞能力；所述舒适性要求用于表征所述目标设备的乘坐舒适度；所述高效性要求用于表征所述目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；

处理模块 220，用于基于所述 AI 服务标识和 AI 服务质量要求，向目标设备反馈第一服务的服务结果。

可选地，所述指示信息通过以下至少一项表示所述 AI 服务质量要求：

质量等级、质量要求指标门限；其中，所述质量要求指标门限包括以下至少一项：安全性要求指标门限、舒适性要求指标门限、高效性要求指标门限和时延要求指标门限。

可选地，接收模块 210，还用于：

获取用于确定质量评估方式的信息；所述质量评估方式用于对服务结果的质量进行评估；所述用于确定质量评估方式的信息包括：质量评估方式对应的可执行文件和运行信息。

可选地，所述运行信息包括以下至少一项：

运行环境配置信息、文件名称、输入格式、输出格式、指示预处理方式的信息、指示后处理方式的信息。

可选地，所述接收模块 210，具体用于执行以下至少一项：

接收来自所述目标设备的所述用于确定质量评估方式的信息；

接收来自所述目标设备的第三方链接，并基于所述第三方链接获取所述用于确定质量评估方式的信息；

接收来自所述目标设备的应用程序编程接口 API，并基于所述 API 获取所述用于确定质量评估方式的信息。

可选地，在所述 AI 服务质量要求包括所述安全性要求的情况下，所述第一信息包括所述安全性要求指标门限和用于评估安全性质量的第一参数信息；其中，第一参数信息包括以下至少一项：

第一指示信息、第二信息、第一加权信息；

其中，第一指示信息用于指示舒适性质量评估包括的至少一个第一目标参数项；

所述第一加权信息包括各个所述第一目标参数项的加权值；

所述第一目标参数项包括以下至少一项：

第一距离、第一保护间隔门限、第二保护间隔门限；

其中，第一距离是所述目标设备与障碍物的距离；所述第一保护间隔门限用于表示最高安全等级的保护间隔下限，所述第二保护间隔门限用于表示最低安全等级的保护间隔下限；

第二信息为各个所述第一目标参数项与安全代价的函数映射信息，或者第一距离与安全等级的对应关系。

可选地，在所述第一距离位于第一保护间隔门限和第二保护间隔门限之间时，安全等级越高，所述第一距离越大。

可选地，所述处理模块 220 还用于：

根据以下至少一项信息，确定所述第一距离；

所述至少一项信息包括：目标设备的尺寸、航向角、目标设备的位置和障碍物边

界。

可选地，在所述 AI 服务质量要求包括所述舒适性要求的情况下，所述第一信息包括所述舒适性要求指标门限和用于评估舒适性质量的第二参数信息；其中，所述第二参数信息包括以下至少一项：

5 第二指示信息、第三信息、第二加权信息；

其中，第二指示信息用于指示舒适性质量评估包括的至少一个第二目标参数项，第三信息为各个所述第二目标参数项对应的舒适性代价函数的信息；所述第二加权信息包括各个所述第二目标参数项的加权值；所述第二目标参数项包括以下至少一项：

10 路径的朝向偏差、路径的航向角、路径的曲率、路径的曲率导数、路径的曲率变化率、加速度、加加速度。

可选地，在所述 AI 服务质量要求包括所述高效性要求的情况下，所述第一信息包括所述高效性要求指标门限和用于高效性质量评估的第三参数信息；其中，所述第三参数信息包括以下至少一项：

第三指示信息、第四信息、第三加权信息；

15 其中，第三指示信息用于指示高效性质量评估包括的至少一个第三目标参数项，第四信息为各个所述第三目标参数项对应的高效性代价函数的信息；所述第三加权信息包括各个所述第三目标参数项的加权值；所述第三目标参数项包括以下至少一项：

与车道中心线的偏差、与理想无障碍路径的偏差、超过道路边界的惩罚值、路径行驶时长、路径行驶距离。

20 可选地，所述 AI 服务质量要求还包括以下至少一种类型的要求：

速度保持代价；

路径单调性；

目标设备速度上限；

路径规划的总时长；

25 路径规划的时间间隔；

其中，速度保持代价为基于服务结果中包括的行驶速度与目标行驶速度的差值得到的，目标行驶速度为由第五信息确定的；所述第五信息包括以下至少一项：路径的曲率、交通规则；

30 路径单调性是指在一次规划的路径中所述目标设备与路径起点的距离是单调递增的，或所述目标设备与路径终点的距离是单调递减的；

目标设备速度上限为基于交通规则和/或目标设备动力学得到的。

可选地，所述第一信息中还包括：用于计算速度保持代价的信息，所述用于计算速度保持代价的信息包括以下至少一项：

偏离速度惩罚函数信息；

35 偏离速度的惩罚倍数。

可选地，所述 AI 服务质量要求还包括：用于指示服务质量要求类型的指示信息，以及每个类型的服务质量要求的合并方式，所述合并方式包括以下至少一项：

每个类型的服务质量要求包括的参数项合并的加权函数信息；

每个类型的服务质量要求包括的参数项的加权值。

40 可选地，第一信息中还包括以下至少一项：

目标设备的尺寸；

目标设备的位置；

目标设备对周围物体的感知结果；所述感知结果包括以下至少一项的感知结果：类型、距离、速度、加速度、边界；

目标设备对周围物体的原始感知数据。

可选地，所述处理模块 220 还用于：

- 5 基于所述第一信息和质量评估方式的信息，或，基于所述第一信息确定第三信息，所述第三信息包括以下至少一项：质量要求指标、质量要求指标计算方法、质量要求指标输入数据和格式、质量要求指标门限。

可选地，所述处理模块 220 还用于：

- 10 基于所述第三信息和所述服务结果，确定所述服务结果的第一评价结果；
所述网络侧设备向所述目标设备发送所述服务结果的第一评价结果。

可选地，所述接收模块 210 还用于：

接收所述目标设备发送的第二评价结果，所述第二评价结果为所述目标设备对所述服务结果进行评价的评价结果。

可选地，第一评价结果或第二评价结果包括以下至少一项：

- 15 AI 服务标识；
AI 服务标识对应的服务等级；
AI 服务标识对应的服务质量数值；
综合服务质量数值；
综合服务质量等级。

20 可选地，所述处理模块 220 具体用于：

基于所述 AI 服务标识和第四信息，确定提供服务的算法信息；所述第四信息包括以下至少一项：AI 服务质量要求、综合服务质量要求；

基于所述 AI 服务标识以及所述算法信息，向目标设备反馈所述服务结果。

- 25 可选地，在提供的 AI 服务为训练服务的情况下，所述算法信息包括训练的奖励函数；

在提供的 AI 服务为推理服务的情况下，所述算法信息包括推理模型。

可选地，所述处理模块 220 还用于：

基于所述 AI 服务标识、推理服务对应的服务结果和综合服务质量要求，确定推理服务对应的服务结果的第一评价结果。

- 30 可选地，所述综合服务质量要求为由所述第一信息确定的，包括以下至少一种方式：

基于所述第一信息直接得到的；

基于所述第一信息得到 AI 服务质量要求，再基于所述 AI 服务质量要求得到的。

- 35 可选地，在所述综合服务质量要求为由所述第一信息确定的情况下，所述第一信息还包括以下至少一项：

综合服务质量等级；
综合服务质量门限；
各服务质量的权重；
各服务质量的加权方法。

- 40 本实施例的装置，可以用于执行前述网络侧方法实施例中任一实施例的方法，其具体实现过程与技术效果与网络侧方法实施例中类似，具体可以参见网络侧方法实施例中的详细介绍，此处不再赘述。

图 6 是本申请提供的 AI 服务的处理装置的结构示意图之二。如图 6 所示，本实施例提供的 AI 服务的处理装置，包括：

发送模块 310，用于向网络侧设备发送第一信息，所述第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；所述 AI 服务质量要求包括以下至少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；所述安全性要求用于表征所述目标设备的防碰撞能力；所述舒适性要求用于表征所述目标设备的乘坐舒适度；所述高效性要求用于表征所述目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；

接收模块 320，用于接收所述网络侧设备发送的第一服务的服务结果。

可选地，所述指示信息通过以下至少一项表示所述 AI 服务质量要求：

质量等级、质量要求指标门限；其中，所述质量要求指标门限包括以下至少一项：安全性要求指标门限、舒适性要求指标门限、高效性要求指标门限和时延要求指标门限。

可选地，所述发送模块 310，还用于：

向所述网络侧设备发送用于确定质量评估方式的信息；所述质量评估方式用于对服务结果的质量进行评估；所述用于确定质量评估方式的信息包括：质量评估方式对应的可执行文件和运行信息。

可选地，所述运行信息包括以下至少一项：

运行环境配置信息、文件名称、输入格式、输出格式、指示预处理方式的信息、指示后处理方式的信息。

可选地，所述发送模块 310，具体用于执行以下至少一项：

向所述网络侧设备发送第三方链接，所述第三方链接用于获取所述用于确定质量评估方式的信息；

向所述网络侧设备发送应用程序编程接口 API，所述应用程序编程接口 API 用于获取所述用于确定质量评估方式的信息。

可选地，在所述 AI 服务质量要求包括所述安全性要求的情况下，所述第一信息包括所述安全性要求指标门限和用于评估安全性质量的第一参数信息；其中，第一参数信息包括以下至少一项：

第一指示信息、第二信息、第一加权信息；

其中，第一指示信息用于指示舒适性质量评估包括的至少一个第一目标参数项；所述第一加权信息包括各个所述第一目标参数项的加权值；

所述第一目标参数项包括以下至少一项：

第一距离、第一保护间隔门限、第二保护间隔门限；

其中，第一距离是所述目标设备与障碍物的距离；所述第一保护间隔门限用于表示最高安全等级的保护间隔下限，所述第二保护间隔门限用于表示最低安全等级的保护间隔下限；

第二信息为各个所述第一目标参数项与安全代价的函数映射信息，或者第一距离与安全等级的对应关系。

可选地，在所述第一距离位于第一保护间隔门限和第二保护间隔门限之间时，安全等级越高，所述第一距离越大。

可选地，所述第一距离为基于以下至少一项信息确定的；

所述至少一项信息包括：目标设备的尺寸、航向角、目标设备的位置和障碍物边

界。

可选地，在所述 AI 服务质量要求包括所述舒适性要求的情况下，所述第一信息包括所述舒适性要求指标门限和用于评估舒适性质量的第二参数信息；其中，所述第二参数信息包括以下至少一项：

5 第二指示信息、第三信息、第二加权信息；

其中，第二指示信息用于指示舒适性质量评估包括的至少一个第二目标参数项，第三信息为各个所述第二目标参数项对应的舒适性代价函数的信息；所述第二加权信息包括各个所述第二目标参数项的加权值；所述第二目标参数项包括以下至少一项：

10 路径的朝向偏差、路径的航向角、路径的曲率、路径的曲率导数、路径的曲率变化率、加速度、加加速度。

可选地，在所述 AI 服务质量要求包括所述高效性要求的情况下，所述第一信息包括所述高效性要求指标门限和用于高效性质量评估的第三参数信息；其中，所述第三参数信息包括以下至少一项：

第三指示信息、第四信息、第三加权信息；

15 其中，第三指示信息用于指示高效性质量评估包括的至少一个第三目标参数项，第四信息为各个所述第三目标参数项对应的高效性代价函数的信息；所述第三加权信息包括各个所述第三目标参数项的加权值；所述第三目标参数项包括以下至少一项：

与车道中心线的偏差、与理想无障碍路径的偏差、超过道路边界的惩罚值、路径行驶时长、路径行驶距离。

20 可选地，所述 AI 服务质量要求还包括以下至少一种类型的要求：

速度保持代价；

路径单调性；

目标设备速度上限；

路径规划的总时长；

25 路径规划的时间间隔；

其中，速度保持代价为基于服务结果中包括的行驶速度与目标行驶速度的差值得到的，目标行驶速度为由第五信息确定的；所述第五信息包括以下至少一项：路径的曲率、交通规则；

30 路径单调性是指在一次规划的路径中所述目标设备与路径起点的距离是单调递增的，或所述目标设备与路径终点的距离是单调递减的；

目标设备速度上限为基于交通规则和/或目标设备动力学得到的。

可选地，所述第一信息中还包括：用于计算速度保持代价的信息，所述用于计算速度保持代价的信息包括以下至少一项：

偏离速度惩罚函数信息；

35 偏离速度的惩罚倍数。

可选地，所述 AI 服务质量要求还包括用于指示服务质量要求类型的指示信息，以及每个类型的服务质量要求的合并方式，所述合并方式包括以下至少一项：

每个类型的服务质量要求包括的参数项合并的加权函数信息；

每个类型的服务质量要求包括的参数项的加权值。

40 可选地，第一信息中还包括以下至少一项：

目标设备的尺寸；

目标设备的位置；

目标设备对周围物体的感知结果；所述感知结果包括以下至少一项的感知结果：类型、距离、速度、加速度、边界；

目标设备对周围物体的原始感知数据。

可选地，所述接收模块 320 还用于：

- 5 接收所述网络侧设备发送的所述服务结果的第一评价结果；所述服务结果的第一评价结果为基于所述第六信息和所述服务结果确定的；所述第六信息包括以下至少一项：质量要求指标、质量要求指标计算方法、质量要求指标输入数据和格式、质量要求指标门限。

可选地，所述发送模块 310 还用于：

- 10 向所述网络侧设备发送第二评价结果，所述第二评价结果为所述目标设备对所述服务结果进行评价的评价结果。

可选地，第一评价结果或第二评价结果包括以下至少一项：

AI 服务标识；

AI 服务标识对应的服务等级；

- 15 AI 服务标识对应的服务质量数值；

综合服务质量数值；

综合服务质量等级。

可选地，综合服务质量要求为由所述第一信息确定的，包括以下至少一种方式：

基于所述第一信息直接得到的；

- 20 基于所述第一信息得到 AI 服务质量要求，再基于所述 AI 服务质量要求得到的。

可选地，在所述综合服务质量要求为由所述第一信息确定的情况下，所述第一信息还包括以下至少一项：

综合服务质量等级；

综合服务质量门限；

- 25 各服务质量的权重；

各服务质量的加权方法。

本实施例的装置，可以用于执行前述目标设备侧方法实施例中任一实施例的方法，其具体实现过程与技术效果与目标设备侧方法实施例中类似，具体可以参见目标设备侧方法实施例中的详细介绍，此处不再赘述。

- 30 本申请实施例中的 AI 服务的处理装置可以是电子设备，例如具有操作系统的电子设备，也可以是电子设备中的部件，例如集成电路或芯片。该电子设备可以是终端，也可以为除终端之外的其他设备。示例性的，终端可以包括但不限于上述所列举的终端 11 的类型，其他设备可以为服务器、网络附属存储器（Network Attached Storage，NAS）等，本申请实施例不作具体限定。

- 35 本申请实施例提供的 AI 服务的处理装置能够实现图 2 至图 4 的方法实施例实现的各个过程，并达到相同的技术效果，为避免重复，这里不再赘述。

- 40 可选地，如图 7 所示，本申请实施例还提供一种通信设备 700，包括处理器 701 和存储器 702，存储器 702 上存储有可在所述处理器 701 上运行的程序或指令，例如，该通信设备 700 为目标设备时，该程序或指令被处理器 701 执行时实现上述 AI 服务的处理方法实施例的各个步骤，且能达到相同的技术效果。该通信设备 700 为网络侧设备时，该程序或指令被处理器 701 执行时实现上述 AI 服务的处理方法实施例的各个步骤，且能达到相同的技术效果，为避免重复，这里不再赘述。

本申请实施例还提供一种目标设备，包括处理器和通信接口，所述通信接口用于向网络侧设备发送第一信息，所述第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；所述 AI 服务质量要求包括以下至少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；所述安全性要求用于表征所述目标设备的防碰撞能力；所述舒适性要求用于表征所述目标设备的乘坐舒适度；所述高效性要求用于表征所述目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；所述通信接口还用于接收所述网络侧设备发送的第一服务的服务结果。该目标设备实施例与上述目标设备侧方法实施例对应，上述方法实施例的各个实施过程和实现方式均可适用于该终端实施例中，且能达到相同的技术效果。目标设备具体地，图 8 为实现本申请实施例的一种终端的硬件结构示意图。

该终端 1000 包括但不限于：射频单元 1001、网络模块 1002、音频输出单元 1003、输入单元 1004、传感器 1005、显示单元 1006、用户输入单元 1007、接口单元 1008、存储器 1009、以及处理器 1010 等中的至少部分部件。

本领域技术人员可以理解，终端 1000 还可以包括给各个部件供电的电源（比如电池），电源可以通过电源管理系统与处理器 1010 逻辑相连，从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。图 8 中示出的终端结构并不构成对终端的限定，终端可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件布置，在此不再赘述。

应理解的是，本申请实施例中，输入单元 1004 可以包括图形处理单元（Graphics Processing Unit, GPU）10041 和麦克风 10042，图形处理器 10041 对在视频捕获模式或图像捕获模式中由图像捕获装置（如摄像头）获得的静态图片或视频的图像数据进行处理。显示单元 1006 可包括显示面板 10061，可以采用液晶显示器、有机发光二极管等形式来配置显示面板 10061。用户输入单元 1007 包括触控面板 10071 以及其它输入设备 10072 中的至少一种。触控面板 10071，也称为触摸屏。触控面板 10071 可包括触摸检测装置和触摸控制器两个部分。其它输入设备 10072 可以包括但不限于物理键盘、功能键（比如音量控制按键、开关按键等）、轨迹球、鼠标、操作杆，在此不再赘述。

本申请实施例中，射频单元 1001 将接收来自网络侧设备的下行数据接收后，可以传输给处理器 1010 进行处理；另外，射频单元 1001 可以将上行的数据发送给向网络侧设备发送上行数据。通常，射频单元 1001 包括但不限于天线、至少一个放大器、收发信机、耦合器、低噪声放大器、双工器等。

存储器 1009 可用于存储软件程序或指令以及各种数据。存储器 1009 可主要包括存储程序或指令的第一存储区和存储数据的第二存储区，其中，第一存储程序或指令区可存储操作系统、至少一个功能所需的应用程序或指令（比如声音播放功能、图像播放功能等）等。此外，存储器 1009 可以包括易失性存储器或非易失性存储器，或者，存储器 1009 可以包括易失性和非易失性存储器两者。包括高速随机存取存储器，还可以包括非易失性存储器，其中，非易失性存储器可以是只读存储器(Read-Only Memory, ROM)、可编程只读存储器(Programmable ROM, PROM)、可擦除可编程只读存储器(Erasable PROM, EPROM)、电可擦除可编程只读存储器(Electrically EPROM, EEPROM)或闪存。易失性存储器可以是随机存取存储器(Random Access Memory, RAM)，静态随机存取存储器(Static RAM, SRAM)、动态随机存取存储器(Dynamic

RAM, DRAM)、同步动态随机存取存储器(Synchronous DRAM, SDRAM)、双倍数据速率同步动态随机存取存储器(Double Data Rate SDRAM, DDRSDRAM)、增强型同步动态随机存取存储器(Enhanced SDRAM, ESDRAM)、同步连接动态随机存取存储器(Synch link DRAM, SLDRAM)和直接内存总线随机存取存储器(Direct Rambus RAM, DRRAM)。本申请实施例中的存储器 1009 包括但不限于这些和任意其它适合类型的存储器例如至少一个磁盘存储器件、闪存器件、或其它非易失性固态存储器件。

处理器 1010 可包括一个或多个处理单元;可选的,处理器 1010 可集成应用处理器和调制解调处理器,其中,应用处理器主要处理涉及操作系统、用户界面和应用程序或指令等的操作,调制解调处理器主要处理无线通信信号,如基带处理器。可以理解的是,上述调制解调处理器也可以不集成到处理器 1010 中。

其中,射频单元 1001,用于向网络侧设备发送第一信息,所述第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息;所述 AI 服务质量要求包括以下至少一种类型的要求:安全性要求、舒适性要求、高效性要求和时延要求;所述安全性要求用于表征所述目标设备的防碰撞能力;所述舒适性要求用于表征所述目标设备的乘坐舒适度;所述高效性要求用于表征所述目标设备完成任务的效率;所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差;射频单元 1001,还用于接收所述网络侧设备发送的第一服务的服务结果。

可选地,所述指示信息通过以下至少一项表示所述 AI 服务质量要求:

质量等级、质量要求指标门限;其中,所述质量要求指标门限包括以下至少一项:安全性要求指标门限、舒适性要求指标门限、高效性要求指标门限和时延要求指标门限。

可选地,所述射频单元 1001,还用于:

向所述网络侧设备发送用于确定质量评估方式的信息;所述质量评估方式用于对服务结果的质量进行评估;所述用于确定质量评估方式的信息包括:质量评估方式对应的可执行文件和运行信息。

可选地,所述运行信息包括以下至少一项:

运行环境配置信息、文件名称、输入格式、输出格式、指示预处理方式的信息、指示后处理方式的信息。

可选地,所述射频单元 1001,具体用于执行以下至少一项:

向所述网络侧设备发送第三方链接,所述第三方链接用于获取所述用于确定质量评估方式的信息;

向所述网络侧设备发送应用程序编程接口 API,所述应用程序编程接口 API 用于获取所述用于确定质量评估方式的信息。

可选地,在所述 AI 服务质量要求包括所述安全性要求的情况下,所述第一信息包括所述安全性要求指标门限和用于评估安全性质量的第一参数信息;其中,第一参数信息包括以下至少一项:

第一指示信息、第二信息、第一加权信息;

其中,第一指示信息用于指示舒适性质量评估包括的至少一个第一目标参数项;所述第一加权信息包括各个所述第一目标参数项的加权值;

所述第一目标参数项包括以下至少一项:

第一距离、第一保护间隔门限、第二保护间隔门限;

其中,第一距离是所述目标设备与障碍物的距离;所述第一保护间隔门限用于表

示最高安全等级的保护间隔下限，所述第二保护间隔门限用于表示最低安全等级的保护间隔下限；

第二信息为各个所述第一目标参数项与安全代价的函数映射信息，或者第一距离与安全等级的对应关系。

5 可选地，在所述第一距离位于第一保护间隔门限和第二保护间隔门限之间时，安全等级越高，所述第一距离越大。

可选地，所述第一距离为基于以下至少一项信息确定的；

所述至少一项信息包括：目标设备的尺寸、航向角、目标设备的位置和障碍物边界。

10 可选地，在所述 AI 服务质量要求包括所述舒适性要求的情况下，所述第一信息包括所述舒适性要求指标门限和用于评估舒适性质量的第二参数信息；其中，所述第二参数信息包括以下至少一项：

第二指示信息、第三信息、第二加权信息；

其中，第二指示信息用于指示舒适性质量评估包括的至少一个第二目标参数项，

15 第三信息为各个所述第二目标参数项对应的舒适性代价函数的信息；所述第二加权信息包括各个所述第二目标参数项的加权值；所述第二目标参数项包括以下至少一项：

路径的朝向偏差、路径的航向角、路径的曲率、路径的曲率导数、路径的曲率变化率、加速度、加加速度。

20 可选地，在所述 AI 服务质量要求包括所述高效性要求的情况下，所述第一信息包括所述高效性要求指标门限和用于评估高效性质量的第三参数信息；其中，所述第三参数信息包括以下至少一项：

第三指示信息、第四信息、第三加权信息；

其中，第三指示信息用于指示高效性质量评估包括的至少一个第三目标参数项，

25 第四信息为各个所述第三目标参数项对应的高效性代价函数的信息；所述第三加权信息包括各个所述第三目标参数项的加权值；所述第三目标参数项包括以下至少一项：

与车道中心线的偏差、与理想无障碍路径的偏差、超过道路边界的惩罚值、路径行驶时长、路径行驶距离。

可选地，所述 AI 服务质量要求还包括以下至少一种类型的要求：

速度保持代价；

30 路径单调性；

目标设备速度上限；

路径规划的总时长；

路径规划的时间间隔；

35 其中，速度保持代价为基于服务结果中包括的行驶速度与目标行驶速度的差值得到的，目标行驶速度为由第五信息确定的；所述第五信息包括以下至少一项：路径的曲率、交通规则；

路径单调性是指在一次规划的路径中所述目标设备与路径起点的距离是单调递增的，或所述目标设备与路径终点的距离是单调递减的；

目标设备速度上限为基于交通规则和/或目标设备动力学得到的。

40 可选地，所述第一信息中还包括：用于计算速度保持代价的信息，所述用于计算速度保持代价的信息包括以下至少一项：

偏离速度惩罚函数信息；

偏离速度的惩罚倍数。

可选地，所述 AI 服务质量要求还包括用于指示服务质量要求类型的指示信息，以及每个类型的服务质量要求的合并方式，所述合并方式包括以下至少一项：

每个类型的服务质量要求包括的参数项合并的加权函数信息；

5 每个类型的服务质量要求包括的参数项的加权值。

可选地，第一信息中还包括以下至少一项：

目标设备的尺寸；

目标设备的位置；

目标设备对周围物体的感知结果；所述感知结果包括以下至少一项的感知结果：

10 类型、距离、速度、加速度、边界；

目标设备对周围物体的原始感知数据。

可选地，所述射频单元 1001 还用于：

接收所述网络侧设备发送的所述服务结果的第一评价结果；所述服务结果的第一评价结果为基于所述第六信息和所述服务结果确定的；所述第六信息包括以下至少一项：质量要求指标、质量要求指标计算方法、质量要求指标输入数据和格式、质量要求指标门限。

可选地，所述射频单元 1001 还用于：

向所述网络侧设备发送第二评价结果，所述第二评价结果为所述目标设备对所述服务结果进行评价的评价结果。

20 可选地，第一评价结果或第二评价结果包括以下至少一项：

AI 服务标识；

AI 服务标识对应的服务等级；

AI 服务标识对应的服务质量数值；

综合服务质量数值；

25 综合服务质量等级。

可选地，综合服务质量要求为由所述第一信息确定的，包括以下至少一种方式：

基于所述第一信息直接得到的；

基于所述第一信息得到 AI 服务质量要求，再基于所述 AI 服务质量要求得到的。

30 可选地，在所述综合服务质量要求为由所述第一信息确定的情况下，所述第一信息还包括以下至少一项：

综合服务质量等级；

综合服务质量门限；

各服务质量的权重；

各服务质量的加权方法。

35 本申请实施例还提供一种网络侧设备，包括处理器和通信接口，所述通信接口用于接收来自目标设备的第一信息，所述第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；所述 AI 服务质量要求包括以下至少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；所述安全性要求用于表征所述目标设备的防碰撞能力；所述舒适性要求用于表征所述目标设备的乘坐舒适度；所述高
40 效性要求用于表征所述目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；所述处理器用于基于所述 AI 服务标识和 AI 服务质量要求，向目标设备反馈服务结果。该网络侧设备实施例与上述网

络侧设备方法实施例对应，上述方法实施例的各个实施过程和实现方式均可适用于该网络侧设备实施例中，且能达到相同的技术效果。

具体地，本申请实施例还提供了一种网络侧设备。如图 9 所示，该网络侧设备 800 包括：天线 81、射频装置 82、基带装置 83、处理器 85 和存储器 85。

5 天线 81 与射频装置 82 连接。

在上行方向上，射频装置 82 通过天线 81 接收信息，将接收的信息发送给基带装置 83 进行处理。

在下行方向上，基带装置 83 对要发送的信息进行处理，并发送给射频装置 82，射频装置 82 对收到的信息进行处理后经过天线 81 发送出去。

10 上述频带处理装置可以位于基带装置 83 中，以上实施例中接入网设备执行的方法可以在基带装置 83 中实现，该基带装置 83 包括基带处理器 85 和存储器 85。

基带装置 83 例如可以包括至少一个基带板，该基带板上设置有多个芯片，如图 9 所示，其中一个芯片例如为基带处理器 85，通过总线接口与存储器 85 连接，以调用存储器 85 中的程序，执行以上方法实施例中所示的接入网设备操作。

15 接入网设备 800 还可以包括网络接口 86，用于与射频装置 82 交互信息，该接口例如为通用公共无线接口（common public radio interface，简称 CPRI）。

具体地，本申请实施的接入网设备 800 还包括：存储在存储器 85 上并可在处理器 85 上运行的指令或程序，处理器 85 调用存储器 85 中的指令或程序执行如图 5 所示模块执行的方法，并达到相同的技术效果，为避免重复，故不在此赘述。

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

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

25 本申请实施例另提供了一种芯片，所述芯片包括处理器和通信接口，所述通信接口和所述处理器耦合，所述处理器用于运行程序或指令，实现上述 AI 服务的处理方法实施例的各个过程，且能达到相同的技术效果，为避免重复，这里不再赘述。

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

本申请实施例另提供了一种计算机程序/程序产品，所述计算机程序/程序产品被存储在存储介质中，所述计算机程序/程序产品被至少一个处理器执行以实现上述 AI 服务的处理方法实施例的各个过程，且能达到相同的技术效果，为避免重复，这里不再赘述。

35 本申请实施例还提供了一种通信系统，包括：目标设备及网络侧设备，所述目标设备可用于执行如上所述的 AI 服务的处理方法的步骤，所述网络侧设备可用于执行如上所述的 AI 服务的处理方法的步骤。

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

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

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

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

权 利 要 求 书

1. 一种人工智能 AI 服务的处理方法，包括：

网络侧设备接收来自目标设备的第一信息，所述第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；所述 AI 服务质量要求包括以下至少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；所述安全性要求用于表征所述目标设备的防碰撞能力；所述舒适性要求用于表征所述目标设备的乘坐舒适度；所述高效性要求用于表征所述目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；

所述网络侧设备基于所述 AI 服务标识和 AI 服务质量要求，向目标设备反馈所述第一服务的服务结果。

2. 根据权利要求 1 所述的方法，其中，

所述指示信息通过以下至少一项表示所述 AI 服务质量要求：

质量等级、质量要求指标门限；其中，所述质量要求指标门限包括以下至少一项：安全性要求指标门限、舒适性要求指标门限、高效性要求指标门限和时延要求指标门限。

3. 根据权利要求 1 或 2 所述的方法，其中，所述方法还包括以下至少一项：

所述网络侧设备接收来自所述目标设备的用于确定质量评估方式的信息；

所述网络侧设备接收来自所述目标设备的第三方链接，并基于所述第三方链接获取用于确定质量评估方式的信息；

所述网络侧设备接收来自所述目标设备的应用程序编程接口 API，并基于所述 API 获取用于确定质量评估方式的信息；

其中，所述质量评估方式用于对服务结果的质量进行评估；所述用于确定质量评估方式的信息包括：质量评估方式对应的可执行文件和运行信息；

所述运行信息包括以下至少一项：

运行环境配置信息、文件名称、输入格式、输出格式、指示预处理方式的信息、指示后处理方式的信息。

4. 根据权利要求 1 或 2 所述的方法，其中，

在所述 AI 服务质量要求包括所述安全性要求的情况下，所述第一信息包括所述安全性要求指标门限和用于评估安全性质量的第一参数信息；其中，第一参数信息包括以下至少一项：

第一指示信息、第二信息、第一加权信息；

其中，第一指示信息用于指示安全性质量评估包括的至少一个第一目标参数项；所述第一加权信息包括各个所述第一目标参数项的加权值；

所述第一目标参数项包括以下至少一项：

第一距离、第一保护间隔门限、第二保护间隔门限；

其中，第一距离是所述目标设备与障碍物的距离；所述第一保护间隔门限用于表示最高安全等级的保护间隔下限，所述第二保护间隔门限用于表示最低安全等级的保护间隔下限；

第二信息为各个所述第一目标参数项与安全代价的函数映射信息，或者第一距离与安全等级的对应关系；

所述第一距离为由以下至少一项信息确定的；所述至少一项信息包括：目标设备

的尺寸、航向角、目标设备的位置和障碍物边界。

5. 根据权利要求 1 或 2 所述的方法，其中，

在所述 AI 服务质量要求包括所述舒适性要求的情况下，所述第一信息包括所述舒适性要求指标门限和用于评估舒适性质量的第二参数信息；其中，所述第二参数信息包括以下至少一项：

第二指示信息、第三信息、第二加权信息；

其中，第二指示信息用于指示舒适性质量评估包括的至少一个第二目标参数项，第三信息为各个所述第二目标参数项对应的舒适性代价函数的信息；所述第二加权信息包括各个所述第二目标参数项的加权值；所述第二目标参数项包括以下至少一项：

10 路径的朝向偏差、路径的航向角、路径的曲率、路径的曲率导数、路径的曲率变化率、加速度、加加速度。

6. 根据权利要求 1 或 2 所述的方法，其中，

在所述 AI 服务质量要求包括所述高效性要求的情况下，所述第一信息包括所述高效性要求指标门限和用于评估高效性质量的第三参数信息；其中，所述第三参数信息包括以下至少一项：

第三指示信息、第四信息、第三加权信息；

其中，第三指示信息用于指示高效性质量评估包括的至少一个第三目标参数项，第四信息为各个所述第三目标参数项对应的高效性代价函数的信息；所述第三加权信息包括各个所述第三目标参数项的加权值；所述第三目标参数项包括以下至少一项：

20 与车道中心线的偏差、与理想无障碍路径的偏差、超过道路边界的惩罚值、路径行驶时长、路径行驶距离。

7. 根据权利要求 1 或 2 所述的方法，其中，

所述 AI 服务质量要求还包括以下至少一种类型的要求：

速度保持代价；

25 路径单调性；

目标设备速度上限；

路径规划的总时长；

路径规划的时间间隔；

其中，速度保持代价为基于服务结果中包括的行驶速度与目标行驶速度的差值得到的，目标行驶速度为由第五信息确定的；所述第五信息包括以下至少一项：路径的曲率、交通规则；

路径单调性是指在一次规划的路径中所述目标设备与路径起点的距离是单调递增的，或所述目标设备与路径终点的距离是单调递减的；

目标设备速度上限为基于交通规则和/或目标设备动力学得到的；

35 所述第一信息中还包括：用于计算速度保持代价的信息，所述用于计算速度保持代价的信息包括以下至少一项：

偏离速度惩罚函数信息；

偏离速度的惩罚倍数。

8. 根据权利要求 4-6 任一项所述的方法，其中，

40 所述 AI 服务质量要求还包括：用于指示服务质量要求类型的指示信息，以及每个类型的服务质量要求的合并方式，所述合并方式包括以下至少一项：

每个类型的服务质量要求包括的参数项合并的加权函数信息；

每个类型的服务质量要求包括的参数项的加权值。

9. 根据权利要求 1 或 2 所述的方法，其中，

第一信息中还包括以下至少一项：

目标设备的尺寸；

5 目标设备的位置；

目标设备对周围物体的感知结果；所述感知结果包括以下至少一项的感知结果：
类型、距离、速度、加速度、边界；

目标设备对周围物体的原始感知数据。

10. 根据权利要求 1-3 任一项所述的方法，其中，所述方法还包括：

10 所述网络侧设备基于所述第一信息和质量评估方式的信息，或，基于所述第一信息确定第六信息，所述第六信息包括以下至少一项：质量要求指标、质量要求指标计算方法、质量要求指标输入数据和格式、质量要求指标门限。

11. 根据权利要求 10 所述的方法，其中，所述方法还包括：

15 所述网络侧设备基于所述第六信息和所述服务结果，确定所述服务结果的第一评价结果；

所述网络侧设备向所述目标设备发送所述服务结果的第一评价结果；

所述第一评价结果包括以下至少一项：

AI 服务标识；

AI 服务标识对应的服务等级；

20 AI 服务标识对应的服务质量数值；

综合服务质量数值；

综合服务质量等级。

12. 根据权利要求 1 或 2 所述的方法，其中，所述方法还包括：

25 所述网络侧设备接收所述目标设备发送的第二评价结果，所述第二评价结果为所述目标设备对所述服务结果进行评价的评价结果；所述第二评价结果包括以下至少一项：

AI 服务标识；

AI 服务标识对应的服务等级；

AI 服务标识对应的服务质量数值；

30 综合服务质量数值；

综合服务质量等级。

13. 根据权利要求 1-3 任一项所述的方法，其中，所述网络侧设备基于所述 AI 服务标识和 AI 服务质量要求，向目标设备反馈服务结果，包括：

35 所述网络侧设备基于所述 AI 服务标识和第七信息，确定提供服务的算法信息；所述第七信息包括以下至少一项：AI 服务质量要求、综合服务质量要求；

所述网络侧设备基于所述 AI 服务标识以及所述算法信息，向目标设备反馈所述服务结果。

14. 根据权利要求 13 所述的方法，其中，所述综合服务质量要求为由所述第一信息确定的，包括以下至少一种方式：

40 基于所述第一信息直接得到的；

基于所述第一信息得到 AI 服务质量要求，再基于所述 AI 服务质量要求得到的；
所述第一信息还包括以下至少一项：

综合服务质量等级；
综合服务质量门限；
各服务质量的权重；
各服务质量的加权方法。

5 15. 一种 AI 服务的处理方法，包括：

目标设备向网络侧设备发送第一信息，所述第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；所述 AI 服务质量要求包括以下至少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；所述安全性要求用于表征所述目标设备的防碰撞能力；所述舒适性要求用于表征所述目标设备的乘坐舒适度；所述高效性要求用于表征所述目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；

所述目标设备接收所述网络侧设备发送的所述第一服务的服务结果。

16. 根据权利要求 15 所述的方法，其中，

所述指示信息通过以下至少一项表示所述 AI 服务质量要求：

15 质量等级、质量要求指标门限；其中，所述质量要求指标门限包括以下至少一项：安全性要求指标门限、舒适性要求指标门限、高效性要求指标门限和时延要求指标门限。

17. 根据权利要求 15 或 16 所述的方法，其中，所述方法还包括以下至少一项：

所述目标设备向所述网络侧设备发送用于确定质量评估方式的信息；

20 所述目标设备向所述网络侧设备发送第三方链接，所述第三方链接用于获取用于确定质量评估方式的信息；

所述目标设备向所述网络侧设备发送应用程序编程接口 API，所述应用程序编程接口 API 用于获取用于确定质量评估方式的信息；

25 其中，所述质量评估方式用于对服务结果的质量进行评估；所述用于确定质量评估方式的信息包括：质量评估方式对应的可执行文件和运行信息；

所述运行信息包括以下至少一项：

运行环境配置信息、文件名称、输入格式、输出格式、指示预处理方式的信息、指示后处理方式的信息。

18. 根据权利要求 15 或 16 所述的方法，其中，

30 在所述 AI 服务质量要求包括所述安全性要求的情况下，所述第一信息包括所述安全性要求指标门限和用于评估安全性质量的第一参数信息；其中，第一参数信息包括以下至少一项：

第一指示信息、第二信息、第一加权信息；

其中，第一指示信息用于指示安全性质量评估包括的至少一个第一目标参数项；

35 所述第一加权信息包括各个所述第一目标参数项的加权值；

所述第一目标参数项包括以下至少一项：

第一距离、第一保护间隔门限、第二保护间隔门限；

40 其中，第一距离是所述目标设备与障碍物的距离；所述第一保护间隔门限用于表示最高安全等级的保护间隔下限，所述第二保护间隔门限用于表示最低安全等级的保护间隔下限；

第二信息为各个所述第一目标参数项与安全代价的函数映射信息，或者第一距离与安全等级的对应关系；

所述第一距离为基于以下至少一项信息确定的；所述至少一项信息包括：目标设备的尺寸、航向角、目标设备的位置和障碍物边界。

19. 根据权利要求 15 或 16 所述的方法，其中，

5 在所述 AI 服务质量要求包括所述舒适性要求的情况下，所述第一信息包括所述舒适性要求指标门限和用于评估舒适性质量的第二参数信息；其中，所述第二参数信息包括以下至少一项：

第二指示信息、第三信息、第二加权信息；

10 其中，第二指示信息用于指示舒适性质量评估包括的至少一个第二目标参数项，第三信息为各个所述第二目标参数项对应的舒适性代价函数的信息；所述第二加权信息包括各个所述第二目标参数项的加权值；所述第二目标参数项包括以下至少一项：

路径的朝向偏差、路径的航向角、路径的曲率、路径的曲率导数、路径的曲率变化率、加速度、加加速度。

20. 根据权利要求 15 或 16 所述的方法，其中，

15 在所述 AI 服务质量要求包括所述高效性要求的情况下，所述第一信息包括所述高效性要求指标门限和用于评估高效性质量的第三参数信息；其中，所述第三参数信息包括以下至少一项：

第三指示信息、第四信息、第三加权信息；

20 其中，第三指示信息用于指示高效性质量评估包括的至少一个第三目标参数项，第四信息为各个所述第三目标参数项对应的高效性代价函数的信息；所述第三加权信息包括各个所述第三目标参数项的加权值；所述第三目标参数项包括以下至少一项：

与车道中心线的偏差、与理想无障碍路径的偏差、超过道路边界的惩罚值、路径行驶时长、路径行驶距离。

21. 根据权利要求 15 或 16 所述的方法，其中，

25 所述 AI 服务质量要求还包括以下至少一种类型的要求：

速度保持代价；

路径单调性；

目标设备速度上限；

路径规划的总时长；

路径规划的时间间隔；

30 其中，速度保持代价为基于服务结果中包括的行驶速度与目标行驶速度的差值得到的，目标行驶速度为由第五信息确定的；所述第五信息包括以下至少一项：路径的曲率、交通规则；

路径单调性是指在一次规划的路径中所述目标设备与路径起点的距离是单调递增的，或所述目标设备与路径终点的距离是单调递减的；

35 目标设备速度上限为基于交通规则和/或目标设备动力学得到的；

所述第一信息中还包括：用于计算速度保持代价的信息，所述用于计算速度保持代价的信息包括以下至少一项：

偏离速度惩罚函数信息；

偏离速度的惩罚倍数。

40 22. 根据权利要求 18-20 任一项所述的方法，其中，

所述 AI 服务质量要求还包括用于指示服务质量要求类型的指示信息，以及每个类型的服务质量要求的合并方式，所述合并方式包括以下至少一项：

每个类型的服务质量要求包括的参数项合并的加权函数信息；

每个类型的服务质量要求包括的参数项的加权值。

23. 根据权利要求 15 或 16 所述的方法，其中，所述方法还包括：

5 所述目标设备接收所述网络侧设备发送的所述服务结果的第一评价结果；所述服务结果的第一评价结果为基于第六信息和所述服务结果确定的；所述第六信息包括以下至少一项：质量要求指标、质量要求指标计算方法、质量要求指标输入数据和格式、质量要求指标门限；所述第一评价结果包括以下至少一项：

AI 服务标识；

AI 服务标识对应的服务等级；

10 AI 服务标识对应的服务质量数值；

综合服务质量数值；

综合服务质量等级。

24. 根据权利要求 15 或 16 所述的方法，其中，所述方法还包括：

15 所述目标设备向所述网络侧设备发送第二评价结果，所述第二评价结果为所述目标设备对所述服务结果进行评价的评价结果；

所述第二评价结果包括以下至少一项：

AI 服务标识；

AI 服务标识对应的服务等级；

AI 服务标识对应的服务质量数值；

20 综合服务质量数值；

综合服务质量等级。

25. 根据权利要求 15 或 16 所述的方法，其中，综合服务质量要求为由所述第一信息确定的，包括以下至少一种方式：

基于所述第一信息直接得到的；

25 基于所述第一信息得到 AI 服务质量要求，再基于所述 AI 服务质量要求得到的；所述第一信息还包括以下至少一项：

综合服务质量等级；

综合服务质量门限；

各服务质量的权重；

30 各服务质量的加权方法。

26. 一种 AI 服务的处理装置，包括：

接收模块，用于接收来自目标设备的第一信息，所述第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；所述 AI 服务质量要求包括以下至少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；所述安全性要求用于表征评估所述目标设备的防碰撞能力；所述舒适性要求用于表征所述目标设备的乘坐舒适度；所述高效性要求用于表征所述目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；

处理模块，用于基于所述 AI 服务标识和 AI 服务质量要求，向目标设备反馈第一服务的的服务结果。

40 27. 一种 AI 服务的处理装置，包括：

发送模块，用于向网络侧设备发送第一信息，所述第一信息包括第一服务的 AI 服务标识和用于指示 AI 服务质量要求的指示信息；所述 AI 服务质量要求包括以下至

少一种类型的要求：安全性要求、舒适性要求、高效性要求和时延要求；所述安全性要求用于表征目标设备的防碰撞能力；所述舒适性要求用于表征目标设备的乘坐舒适度；所述高效性要求用于表征目标设备完成任务的效率；所述时延要求用于表征所述目标设备发起 AI 服务请求与获取到服务结果之间的时间差；

5 接收模块，用于接收所述网络侧设备发送的第一服务的服务结果。

28. 一种网络侧设备，包括处理器和存储器，所述存储器存储可在所述处理器上运行的程序或指令，所述程序或指令被所述处理器执行时实现如权利要求 1 至 14 任一项所述的 AI 服务的处理方法的步骤。

10 29. 一种目标设备，包括处理器和存储器，所述存储器存储可在所述处理器上运行的程序或指令，所述程序或指令被所述处理器执行时实现如权利要求 15 至 25 任一项所述的 AI 服务的处理方法的步骤。

30. 一种可读存储介质，所述可读存储介质上存储程序或指令，所述程序或指令被处理器执行时实现如权利要求 1 至 14 任一项所述的 AI 服务的处理方法，或者实现如权利要求 15 至 25 任一项所述的 AI 服务的处理方法的步骤。

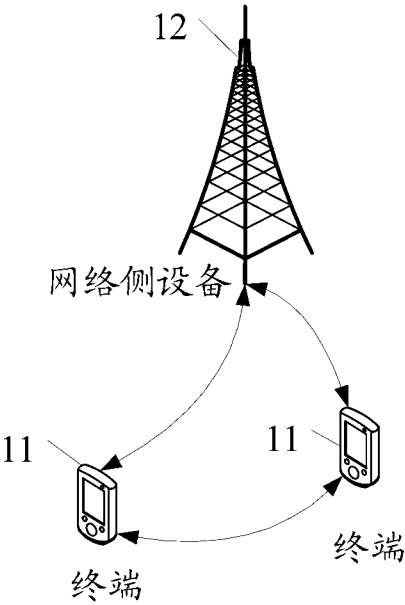


图 1

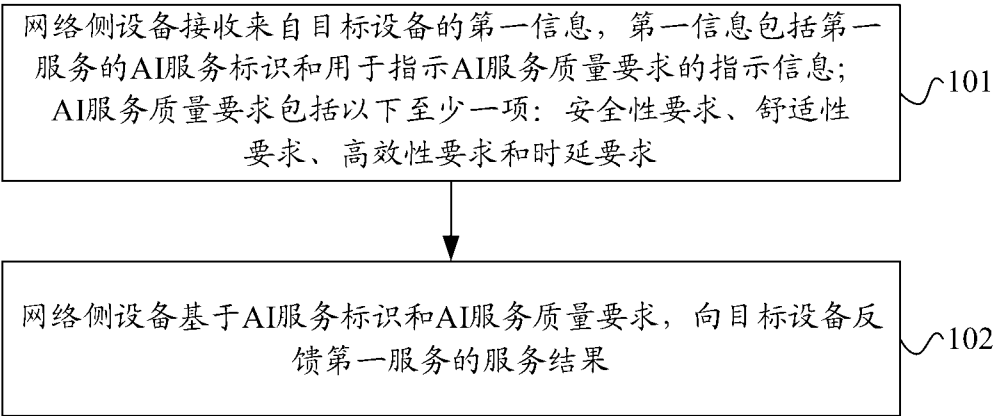


图 2

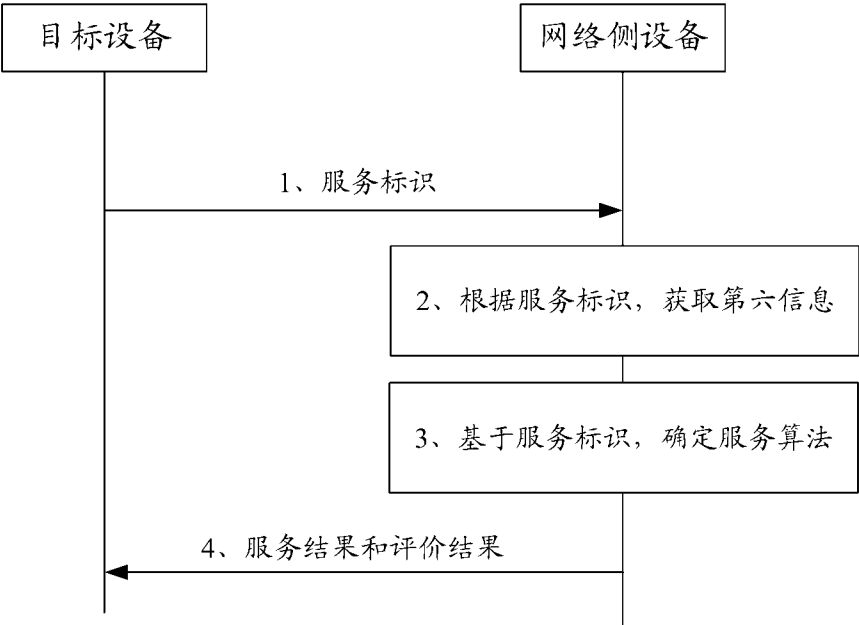


图 3

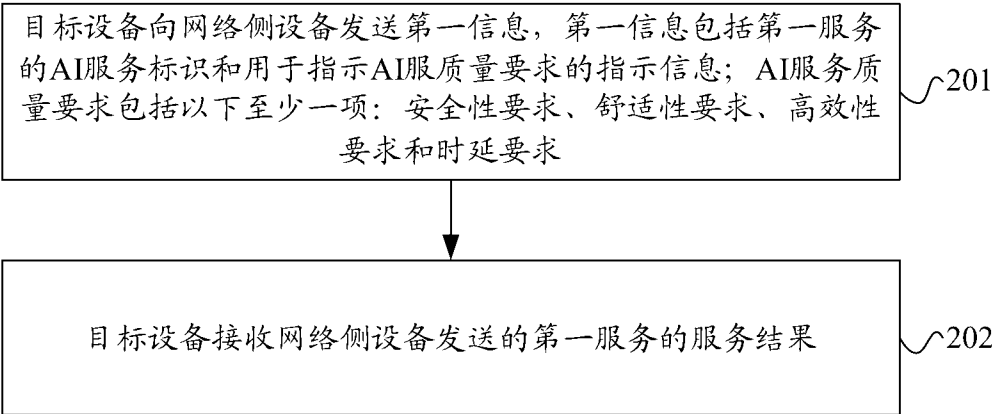


图 4

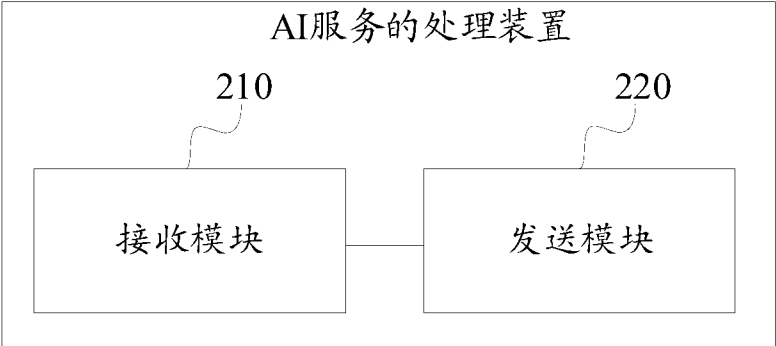


图 5

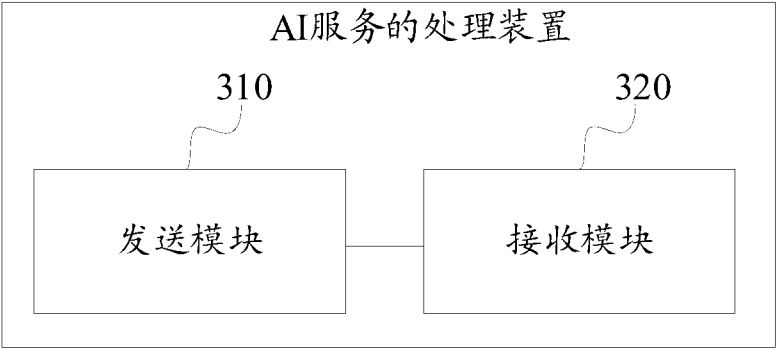


图 6

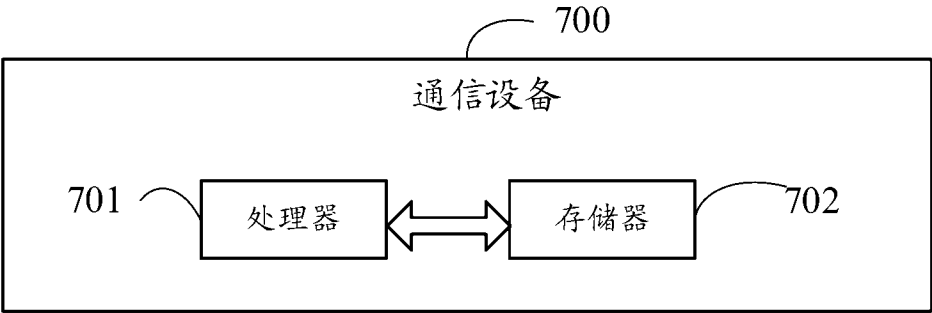


图 7

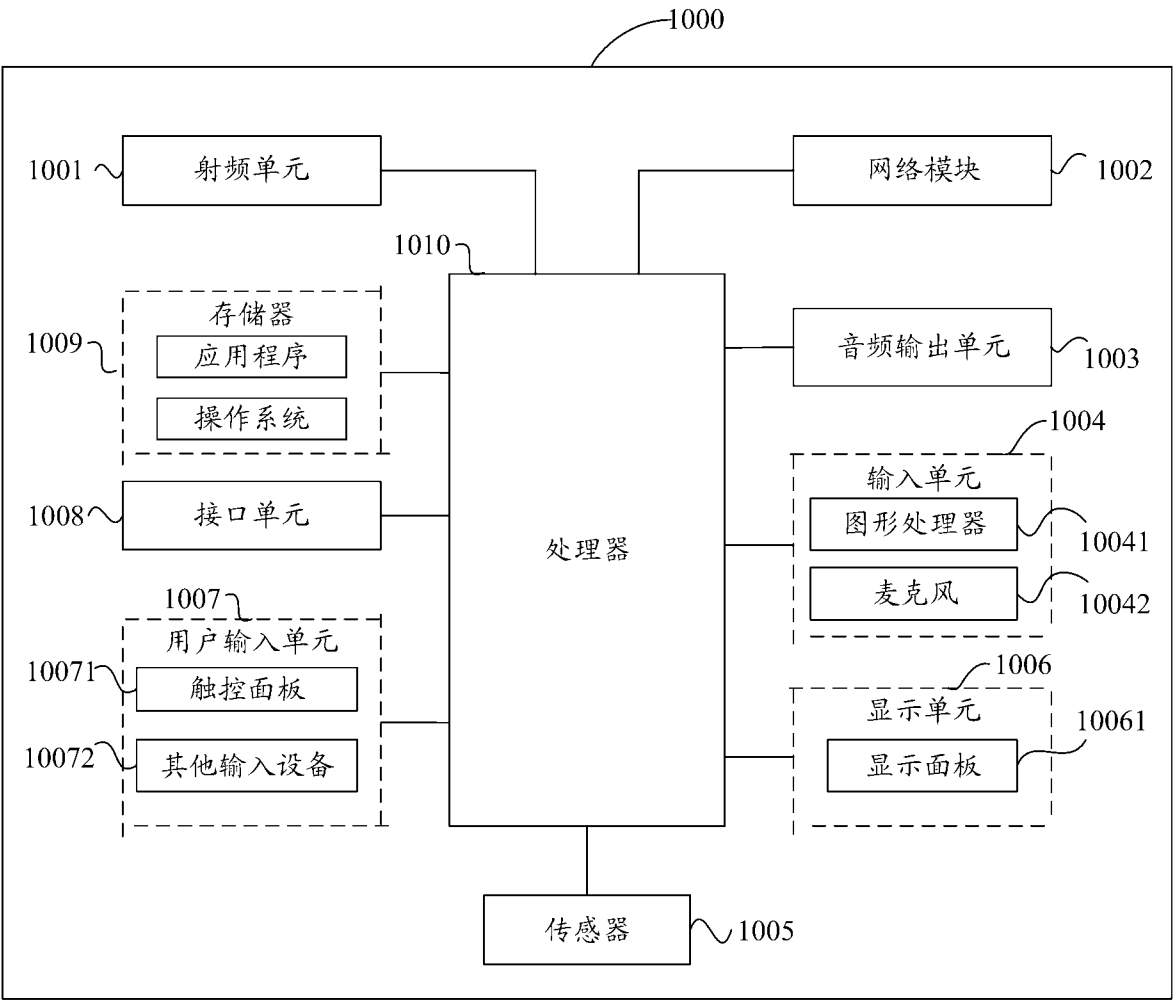


图 8

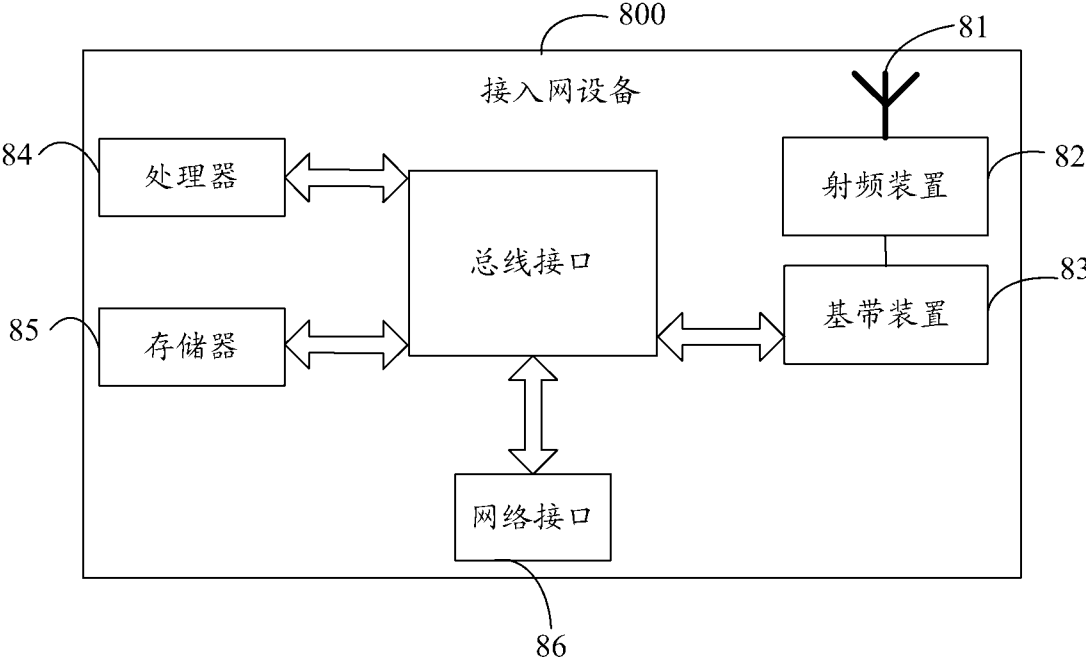


图 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/127307

A. CLASSIFICATION OF SUBJECT MATTER

H04W 4/44(2018.01)i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04W

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; VEN; USTXT; WOTXT; EPTXT; CNKI; 3GPP: 人工智能, 模型, 组件, 标识, 类型, 质量, 服务, 需求, 约束, 部署, 编排, AI, model, service, on-demand, ID, class, user, request, requirement, sequence, QoS, orchestration

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 114745264 A (DATANG MOBILE COMMUNICATIONS EQUIPMENT CO., LTD.) 12 July 2022 (2022-07-12) description, paragraphs [0095]-[0250]	1-3, 10-17, 23-30
Y	CN 114745264 A (DATANG MOBILE COMMUNICATIONS EQUIPMENT CO., LTD.) 12 July 2022 (2022-07-12) description, paragraphs [0095]-[0250]	1-30
Y	CN 114026834 A (INTEL CORP.) 08 February 2022 (2022-02-08) description, paragraphs [0091]-[0163], paragraphs [0348]-[0360], and paragraphs [0496]-[0504], and figure 5	1-30
A	CN 112953778 A (BEIJING JIAOTONG UNIVERSITY) 11 June 2021 (2021-06-11) entire document	1-30
A	CN 114675945 A (TENCENT CLOUD COMPUTING (BEIJING) CO., LTD.) 28 June 2022 (2022-06-28) entire document	1-30



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

02 December 2023

Date of mailing of the international search report

11 December 2023

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/CN2023/127307

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	114745264	A	12 July 2022	None			
CN	114026834	A	08 February 2022	US	2020167196	A1	28 May 2020
				US	11711268	B2	25 July 2023
				US	2022337481	A1	20 October 2022
				DE	102020203877	A1	05 November 2020
				KR	20210149576	A	09 December 2021
				DE	102020203881	A1	05 November 2020
				US	2020167205	A1	28 May 2020
				US	11184236	B2	23 November 2021
				EP	3963866	A2	09 March 2022
				EP	3963866	A4	20 September 2023
				EP	3734453	A1	04 November 2020
				DE	112020000054	T5	11 March 2021
				US	2021144517	A1	13 May 2021
				US	2020136906	A1	30 April 2020
				US	11388054	B2	12 July 2022
				US	2022247635	A1	04 August 2022
				JP	2022530580	A	30 June 2022
				WO	2020226979	A2	12 November 2020
				WO	2020226979	A8	30 December 2020
				WO	2020226979	A3	04 March 2021
CN	112953778	A	11 June 2021	None			
CN	114675945	A	28 June 2022	None			

A. 主题的分类 H04W 4/44(2018.01)i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类		
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) IPC: H04W 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;VEN;USTXT;WOTXT;EPTXT;CNKI;3GPP: 人工智能, 模型, 组件, 标识, 类型, 质量, 服务, 需求, 约束, 部署, 编排, AI, model, service, on-demand, ID, class, user, request, requirement, sequence, QoS, orchestration		
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 114745264 A (大唐移动通信设备有限公司) 2022年7月12日 (2022 - 07 - 12) 说明书第[0095]-[0250]段	1-3、10-17、23-30
Y	CN 114745264 A (大唐移动通信设备有限公司) 2022年7月12日 (2022 - 07 - 12) 说明书第[0095]-[0250]段	1-30
Y	CN 114026834 A (英特尔公司) 2022年2月8日 (2022 - 02 - 08) 说明书第[0091]-[0163]段、第[0348]-[0360]段、第[0496]-[0504]段, 附图5	1-30
A	CN 112953778 A (北京交通大学) 2021年6月11日 (2021 - 06 - 11) 全文	1-30
A	CN 114675945 A (腾讯云计算(北京)有限责任公司) 2022年6月28日 (2022 - 06 - 28) 全文	1-30
☐ 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 2023年12月2日		国际检索报告邮寄日期 2023年12月11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 童雯 电话号码 (+86) 0512-88996096

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2023/127307

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	114745264	A	2022年7月12日	无			
CN	114026834	A	2022年2月8日	US	2020167196	A1	2020年5月28日
				US	11711268	B2	2023年7月25日
				US	2022337481	A1	2022年10月20日
				DE	102020203877	A1	2020年11月5日
				KR	20210149576	A	2021年12月9日
				DE	102020203881	A1	2020年11月5日
				US	2020167205	A1	2020年5月28日
				US	11184236	B2	2021年11月23日
				EP	3963866	A2	2022年3月9日
				EP	3963866	A4	2023年9月20日
				EP	3734453	A1	2020年11月4日
				DE	112020000054	T5	2021年3月11日
				US	2021144517	A1	2021年5月13日
				US	2020136906	A1	2020年4月30日
				US	11388054	B2	2022年7月12日
				US	2022247635	A1	2022年8月4日
				JP	2022530580	A	2022年6月30日
				WO	2020226979	A2	2020年11月12日
				WO	2020226979	A8	2020年12月30日
				WO	2020226979	A3	2021年3月4日
CN	112953778	A	2021年6月11日	无			
CN	114675945	A	2022年6月28日	无			