

一种通信方法及相关设备

技术领域

本申请涉及通信领域，尤其涉及一种通信方法及相关设备。

背景技术

无线通信，可以是两个或两个以上的通信节点间不经由导体或缆线传播而进行的传输通讯，该通信节点一般包括网络设备和终端设备。

目前，在无线通信系统中，通信节点一般具备信号收发能力和计算能力。以具备计算能力的网络设备为例，网络设备的计算能力主要是为信号收发能力提供算力支持（例如：对承载信号的时域资源、频域资源等进行计算），以实现网络设备与其它通信节点的通信。

然而，在通信网络中，通信节点的计算能力除了为上述通信任务提供算力支持之外，还可能具备富余的计算能力。为此，如何利用这些计算能力，是一个亟待解决的技术问题。

发明内容

本申请提供了一种通信方法及相关设备，用于使得通信节点的算力能够应用于人工智能（artificial intelligence，AI）学习，并能够提升不同节点的实现灵活性。

本申请第一方面提供了一种通信方法，该方法由第一节点执行，或者，该方法由第一节点中的部分组件（例如处理器、芯片或芯片系统等）执行，或者该方法还可以由能实现全部或部分第一节点功能的逻辑模块或软件实现。在第一方面及其可能的实现方式中，以该方法由第一节点执行为例进行描述。在该方法中，第一节点接收第一信息，该第一信息指示 N 个配置信息，该 N 个配置信息用于配置 N 个模型参数的资源，N 为正整数；该第一节点基于该 N 个配置信息中的第一配置信息接收第一模型参数，该第一模型参数用于确定与本地信息对应的第一神经网络模型。

基于上述技术方案，第一节点接收的第一信息指示 N 个配置信息，该 N 个配置信息用于配置 N 个模型参数的资源，此后，该第一节点基于该 N 个配置信息中的第一配置信息所配置的资源接收第一模型参数，并且，该第一节点可以基于该第一模型参数确定与本地信息对应的第一神经网络模型。换言之，不同的第一节点均可以基于该第一信息获得与各自的本地信息对应的神经网络模型，后续均可以基于本地获得的神经网络模型参与 AI 学习。相比于不同节点基于随机化的神经网络模型参与 AI 学习的实现方式，在上述技术方案中，不同节点可以基于与各自的本地信息对应的神经网络模型参与 AI 学习。从而，在通信系统中的通信节点作为参与 AI 学习的节点的情况下，能够使得通信节点的算力能够应用于学习系统中的 AI 任务，并能够提升不同节点的实现灵活性。

此外，不同节点的本地信息一般是不同的，而在上述技术方案中，第一节点能够基于第一信息确定与本地信息对应的神经网络模型，为此，不同的第一节点均能够基于与本地信息适配的神经网络模型参与 AI 学习，以提升神经网络模型的处理效率。

可选地，AI 学习可以替换为机器学习，AI 处理，AI 神经网络处理等。

应理解，N 个配置信息所配置的资源包括时域资源、频域资源、空域资源中的一项或多项。

应理解，在 N 个配置信息配置的 N 个模型参数中，模型参数用于神经网络模型的初始化，或者说，模型参数包括神经网络模型的初始化参数。其中，初始化参数可以是用于确定神经网络模型使用的参数，或，初始化参数可以是用于确定神经网络模型训练启动的参数。换言之，对于第一节点而言，该第一节点基于第一配置信息接收第一模型参数之后，该第一节点可以基于该第一模型参数进行神经网络模型的初始化处理，以得到与本地信息对应的第一神经网络模型；或者，该第一节点可以将该第一模型参数作为神经网络模型的初始化参数，完成本地参数初始化之后，得到第一神经网络模型。

在第一方面的一种可能的实现方式中，该本地信息包括本地资源开销水平，本地 AI 能力信息、本地的参数规模中的至少一项。

基于上述技术方案，本地信息可以通过多种方式实现，以提供方案实现的灵活性。

应理解，本地资源开销水平可以理解为本地的计算资源的利用率，本地的存储资源的利用率，本地功

耗的使用情况等。例如，该本地的计算资源可以包括中央处理器（central processor unit, CPU）、微控制器（micro controller unit, MCU）或图形处理器（graphics processing unit, GPU）等模块/器件的资源；该本地的存储资源可以包括只读存储器（read only memory, ROM）、随机存取存储器（random access memory, RAM）等模块/器件的资源；该本地功耗的使用情况可以包括电池模块的健康度，电池模块当前已使用的电量（或电量的百分比），电池模块当前尚未使用的电量（或电量的百分比）等。

应理解，本地 AI 能力信息可以包括本地与 AI 相关的能力信息，例如与 AI 相关的计算能力、存储能力等。类似的，与 AI 相关的计算能力可以包括 CPU、MCU、GPU 等模块/器件的计算能力，与 AI 相关的存储能力可以包括 ROM、RAM 等模块/器件的存储能力。

应理解，本地的参数规模，可以理解为第一节点本地的神经网络模型的参数规模，包括但不限于神经网络模型的神经网络层数、神经网络模型的神经元数量，神经网络模型的输入输出或神经网络模型的计算复杂度等不同。

需要说明的是，第一节点的本地信息可以通过多种方式实现，相应的，对于第一节点而言，该第一节点接收的第一模型参数之后，该第一节点可以通过多种方式确定与本地信息对应的第一神经网络模型。

一种实现示例中，以该本地信息包括本地资源开销水平为例，该本地资源开销水平可以通过 CPU 利用率（或称 CPU 占用率）体现；其中，第一节点可以通过配置或预配置的方式确定不同的 CPU 利用率与不同的神经网络模型之间的映射关系。以不同的神经网络模型包括第一类神经网络模型和第二类神经网络模型为例，第一类神经网络模型可以比第二类神经网络模型更复杂（例如模型参数的规模更大，模型的神经网络层数更多等），该映射关系可以指示 CPU 利用率为 0 至 X%（X 取值为大于 0 且小于 100）的节点对应第一类神经网络模型，并指示 CPU 利用率为 X%至 100%的节点对应第二类神经网络模型。从而，在上述技术方案中，第一节点可以基于本地的 CPU 利用率与 X%进行比较，在本地的 CPU 利用率低于或等于 X%的情况下，第一节点基于第一模型参数得到的第一神经网络模型可以为第一类神经网络模型；在本地 CPU 利用率高于或等于 X%的情况下，第一节点基于第一模型参数得到的第一神经网络模型可以为第二类神经网络模型。

可选地，在上述实现示例中，CPU 利用率的测量可以是在节点执行相同或相似计算参数的这个前提下确定的，即在相对公平的情况下确定 CPU 利用率。比如，针对两个模型的训练，输入参数相同时，不同的节点为了完成一轮训练所需要的 CPU 利用率有差异。例如，在某个节点的 CPU 利用率较低的情况下，可以理解为该节点的 CPU 的算力较强，可以基于模型参数规模较大（或较复杂的模型参数）进行初始化处理；在某个节点的 CPU 利用率较高的情况下，可以理解为该节点的 CPU 的算力较弱，可以基于模型参数规模较小（或较简化的模型参数）进行初始化处理。相应的，基于该差异，可以为该不同的节点配置不同的模型参数，以使得各个节点能够获得差异化（或个性化）的初始化神经网络模型。

另一种实现示例中，以该本地信息包括本地 AI 能力信息为例，该本地 AI 能力信息可以通过 RAM 的存储能力体现；其中，第一节点可以通过配置或预配置的方式确定不同的 RAM 的存储能力与不同的神经网络模型之间的映射关系。以不同的神经网络模型包括第一类神经网络模型、第二类神经网络模型和第三类神经网络模型为例，第一类神经网络模型可以比第二类神经网络模型更复杂（例如模型参数的规模更大，模型的神经网络层数更多等），第二类神经网络模型可以比第三类神经网络模型更复杂；该映射关系可以指示 RAM 的存储能力为 0 至 Y 吉字节（Gigabyte, GB）（Y 的取值为大于 0）的节点对应第一类神经网络模型，并指示 RAM 的存储能力大于 Y GB 且小于或等于 Z GB（Z 的取值为大于 Y）的节点对应第二类神经网络模型，并指示 RAM 的存储能力大于 Z GB 的节点对应第三类神经网络模型。从而，在上述技术方案中，第一节点可以基于本地的 RAM 的存储能力与 Y 和 Z 进行比较，在本地的 RAM 的存储能力低于或等于 Y GB 的情况下，第一节点基于第一模型参数得到的第一神经网络模型可以为第一类神经网络模型；在本地 RAM 的存储能力大于 Y GB 且小于或等于 Z GB 的情况下，第一节点基于第一模型参数得到的第一神经网络模型可以为第二类神经网络模型，在本地 RAM 的存储能力大于 Z GB 的情况下，第一节点基于第一模型参数得到的第一神经网络模型可以为第三类神经网络模型。

在第一方面的一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

基于上述技术方案，N 个配置信息所配置的资源传输的 N 个模型参数包括上述至少一项，以提升方案

实现的灵活性。

在第一方面的一种可能的实现方式中， N 为大于 1 的整数，该 N 个模型参数的参数规模不同，该方法还包括：该第一节点基于该本地信息在该 N 个配置信息中确定该第一配置信息。

基于上述技术方案，在第一信息指示的 N 个配置信息大于 1 的情况下，该 N 个配置信息所配置的 N 个模型参数的参数规模不同，使得第一节点能够基于本地信息在 N 个配置信息中确定与该本地信息对应的第一配置信息，使得该第一节点能够基于该第一配置信息获得与该本地信息对应的模型参数。

应理解， N 个模型参数的参数规模不同，可以理解为该 N 个模型参数对节点的本地信息的要求不同。

例如，以该本地信息包括本地资源开销水平为例， N 个模型参数中参数规模较大的模型参数要求本地资源开销水平低于阈值，以便于提供更多的未被使用的资源处理该参数规模较大的模型参数；又如，以该本地信息包括本地 AI 能力信息为例， N 个模型参数中参数规模较大的模型参数要求本地 AI 能力信息高于阈值，以便于提供更多的 AI 资源处理该参数规模较大的模型参数。

可选地， N 个模型参数的参数规模不同，可以替换为其它的描述，例如 N 个模型参数对应的神经网络层数不同， N 个模型参数对应的神经元不同， N 个模型参数对应的输入输出不同， N 个模型参数的计算复杂度不同，或者， N 个模型参数的计算能力要求不同，或者， N 个模型参数的存储能力要求不同等。

在第一方面的一种可能的实现方式中， N 的取值为 1，其中，该第一模型参数用于配置至少两个第一节点共用的神经网络模型；该方法还包括：该第一节点基于该本地信息将该至少两个第一节点共用的神经网络模型进行知识蒸馏处理，得到该第一神经网络模型。

基于上述技术方案，在第一信息指示的 N 个配置信息为一个配置信息（即第一配置信息）的情况下，该第一配置信息配置的第一模型参数用于配置至少两个第一节点共用的神经网络模型，使得该至少两个第一节点均能够基于该第一配置信息获得共用的神经网络模型，并基于各自的本地信息对该共用的神经网络模型进行知识蒸馏处理，以获得与该本地信息对应的第一神经网络模型。

应理解，至少两个第一节点共用的神经网络模型，可以理解为该至少两个第一节点能够识别/处理的神经网络模型。或者，至少两个第一节点共用的神经网络模型可以表述为：至少两个第一节点通用的神经网络模型，至少两个第一节点能够识别/处理的神经网络模型。

在第一方面的一种可能的实现方式中，该方法还包括：该第一节点接收第二信息，该第二信息指示 M 个配置信息，该 M 个配置信息用于配置 M 个模型参数的传输资源，该 M 个模型参数的参数规模不同， M 为大于 1 的整数；该第一节点基于该 M 个配置信息中的第二配置信息接收第二模型参数，其中，该第二模型参数的参数规模与本地的参数规模相同（或相近，例如该第二模型参数的参数规模与本地的参数规模的差低于阈值）。

基于上述技术方案，第一节点还可以接收指示 M 个配置信息的第二信息，并基于该 M 个配置信息的第二配置信息接收与本地的参数规模相同（或相近）的第二模型参数。换言之，不同的第一节点均可以通过第二信息接收与各自的本地信息的参数规模相同（或相近）的模型参数，并基于所接收的模型参数执行 AI 处理（例如本地训练）。

应理解， M 个配置信息所配置的传输资源用于不同的第一节点接收 M 个模型参数中的其中一个模型参数，使得不同的第一节点后续均可以基于所接收的模型参数执行 AI 处理（例如本地训练）。相应的，该 M 个模型参数可以理解为多个第一节点的全局模型参数，即该 M 个模型参数可以替换为 M 个全局模型参数。

可选地，对于 M 个模型参数中的任一模型参数而言，可以包括梯度和/或权值等。例如，该梯度和/或权值用于初始化的神经网络模型（例如前文描述的第一神经网络模型）的进一步处理，例如神经网络模型的训练，神经网络模型的更新等。

类似地， M 个模型参数的参数规模不同，可以替换为其它的描述，例如 M 个模型参数对应的神经网络层数不同， M 个模型参数对应的神经元不同， M 个模型参数对应的输入输出不同， M 个模型参数的计算复杂度不同，或者， M 个模型参数的计算能力要求不同，或者， M 个模型参数的存储能力要求不同等。

在第一方面的一种可能的实现方式中，该 M 个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的接收功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息。

基于上述技术方案，第二信息指示的 M 个配置信息中的任意一个配置信息可以包括上述至少一项，

5 以提升方案实现的灵活性。

可选地，第一信息指示的 N 个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的接收功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

10 指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息。

在第一方面的一种可能的实现方式中，该方法还包括：该第一节点接收第三信息，该第三信息指示 K 个配置信息，该 K 个配置信息用于配置 K 个模型参数的传输资源，该 K 个模型参数的参数规模不同， K 为大于 1 的整数；该第一节点基于该 K 个配置信息中的第三配置信息发送第三模型参数，其中，该第三模型参数的参数规模与本地的参数规模相同（或相近，例如该第三模型参数的参数规模与本地的参数规模的差低于阈值）。

基于上述技术方案，第一节点还可以接收指示 K 个配置信息的第三信息，并基于该 K 个配置信息的第三配置信息接收与本地的参数规模相同（或相近）的第三模型参数。换言之，不同的第一节点均可以通过第三信息发送与各自的本地信息的参数规模相同（或相近）的模型参数，使得该模型参数的接收方能够对不同的第一节点发送的模型参数执行 AI 处理（例如参数融合处理）。

20 应理解， K 个配置信息所配置的传输资源用于不同的第一节点发送 K 个模型参数中的其中一个模型参数，使得该模型参数的接收方能够对不同的第一节点发送的模型参数执行融合操作。相应的，该 K 个模型参数可以理解为多个第一节点的本地模型参数，即该 K 个模型参数可以替换为 K 个本地模型参数。

可选地，对于 K 个模型参数中的任一模型参数而言，可以包括梯度和/或权值等。例如，该梯度和/或权值用于神经网络模型的进一步融合处理，例如深度融合（Deep Fusion），加权平均融合（Weighted Averaging）等。

类似地， K 个模型参数的参数规模不同，可以替换为其它的描述，例如 K 个模型参数对应的神经网络层数不同， K 个模型参数对应的神经元不同， K 个模型参数对应的输入输出不同， K 个模型参数的计算复杂度不同，或者， K 个模型参数的计算能力要求不同，或者， K 个模型参数的存储能力要求不同等。

30 在第一方面的一种可能的实现方式中，该 K 个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的发射功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

35 指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息；

指示该一个配置信息的传输资源在该 K 个配置信息配置的传输资源中的索引的指示信息。

40 应理解，在该任意一个配置信息包括指示该一个配置信息的传输资源在该 K 个配置信息配置的传输资源中的索引的指示信息的情况下，可以理解为第二节点为该 K 个配置信息配置了传输资源集合（包括时域资源、频域资源、空域资源中的至少一项），使得第一节点能够基于该指示信息的索引在该传输资源集合中确定相应的传输资源。

基于上述技术方案，第三信息指示的 K 个配置信息中的任意一个配置信息可以包括上述至少一项，以提升方案实现的灵活性。

在第一方面的一种可能的实现方式中，在该第一节点接收第一信息之前，该方法还包括：该第一节点接收第四信息，该第四信息用于指示承载请求信息的资源，该请求信息用于请求该 N 个配置信息；该第一节点在该承载请求信息的资源上发送该请求信息。

45 基于上述技术方案，第一节点还可以接收用于指示承载请求信息的资源的第四信息，该请求信息用于

请求该 N 个配置信息，使得该第一节点能够基于该第四信息指示的资源发送请求信息之后，能够通过第一信息获得该 N 个配置信息。

在第一方面的一种可能的实现方式中，该请求信息包括第一正交序列，该第一正交序列对应于该本地信息。

5 可选地，该第一正交序列包含于 P 个正交序列，其中，该 P 个正交序列分别对应于不同的本地信息。换言之，不同第一节点均能够基于各自的本地信息在该 P 个正交序列中确定其中一个正交序列，并向第二节点发送该其中一个正交序列，使得该请求信息的接收方（例如第二节点）能够基于所接收的请求信息获知各个第一节点的本地信息。

10 可选地，在 P 大于 1 的情况下，不同序列可以是正交的关系，即 P 个序列可以表述为 P 个正交序列，以使该 P 个正交序列通过相同资源传输的情况下也能被接收方所识别。而在其它实现方式中，该 P 个序列也可以是通过其它方式被接收方所识别，例如通过频域资源、时域资源的错开。换言之，上述 P 个正交序列可以替换为 P 个序列，第一正交序列可以替换为第一序列。

此外，第二节点作为该请求信息的接收方的情况下，该第二节点能够基于接收的一个或多个正交序列确定各个第一节点的本地信息，可以将各个第一节点的本地信息作为 N 个配置信息以及 M 个配置信息以及 K 个配置信息至少一项的确定依据，使得该第二节点发送的各个配置信息能够与各个第一节点的本地信息相适配。

此外，由于不同的第一节点均能够通过该请求信息指示自身的本地参数规模，能够降低开销。

可选地，第一节点除了可以基于第四信息确定承载该第一正交序列的资源之外，该第一节点也可以通过其它方式（例如预配置的方式）确定承载该第一正交序列的资源，此处不做限定。

20 可选地， P 的取值可以大于或等于 N 的取值。例如，第一节点可以通过配置或预配置的方式获知该 P 个正交序列分别对应的不同参数规模，并且，第二节点可以通过接收来自一个或多个第一节点的请求信息确定存在各个第一节点的参数规模。相应的，该第二节点可以进一步基于所确定的各个第一节点的参数规模生成该 N 个配置信息。一般地， P 个正交序列可以存在冗余设置（即 P 大于或等于 N ），以使得各个第一节点均能够通过请求信息指示各自的本地的参数规模。

25 在第一方面的一种可能的实现方式中，在该第一节点接收第一信息之前，该方法还包括：该第一节点接收第五信息，该第五信息用于指示该 N 个模型参数分别对应的本地信息，和/或，该第五信息用于指示该第一信息是否包括该 N 个配置信息。

30 基于上述技术方案，第一节点还可以接收第五信息，以便于通过该第五信息获知该 N 个配置信息的相关信息，并且基于所获知的信息确定是否接收该 N 个配置信息，能够避免第一节点不必要的检测以降低开销。

在第一方面的一种可能的实现方式中，在该第一节点接收第三信息之前，该方法还包括：该第一节点接收第六信息，该第六信息用于指示该 M 个模型参数分别对应的本地信息，和/或，该第六信息用于指示该第三信息是否包括该 M 个配置信息。

35 基于上述技术方案，第一节点还可以接收第六信息，以便于通过该第六信息获知该 M 个配置信息的相关信息，并且基于所获知的信息确定是否接收该 M 个配置信息，能够避免第一节点不必要的检测以降低开销。

在第一方面的一种可能的实现方式中，在该第一节点接收第二信息之前，该方法还包括：该第一节点接收第七信息，该第七信息用于指示该 K 个模型参数分别对应的本地信息，和/或，该第七信息用于指示该第二信息是否包括该 K 个配置信息。

40 基于上述技术方案，第一节点还可以接收第七信息，以便于通过该第七信息获知该 K 个配置信息的相关信息，并且基于所获知的信息确定是否接收该 K 个配置信息，能够避免第一节点不必要的检测以降低开销。

45 本申请第二方面提供了一种通信方法，该方法由第二节点执行，或者，该方法由第二节点中的部分组件（例如处理器、芯片或芯片系统等）执行，或者该方法还可以由能实现全部或部分第二节点功能的逻辑模块或软件实现。在第二方面及其可能的实现方式中，以该方法由第二节点执行为例进行描述。在该方法

中，第二节点确定第一信息，该第一信息指示 N 个配置信息，该 N 个配置信息用于配置 N 个模型参数的资源， N 为正整数；其中，该 N 个模型参数中的一个模型参数用于确定与第一节点的本地信息对应的神经网络模型；该第二节点发送该第一信息。

基于上述技术方案，第二节点发送的第一信息指示 N 个配置信息，该 N 个配置信息用于配置 N 个模型参数的资源，此后，第一节点在接收该第一信息之后，该第一节点可以基于该 N 个配置信息中的第一配置信息所配置的资源接收第一模型参数，并且，该第一节点可以基于该第一模型参数确定与本地信息对应的第一神经网络模型。换言之，不同的第一节点均可以基于该第一信息获得与各自的本地信息对应的神经网络模型，后续均可以基于本地获得的神经网络模型参与 AI 学习。相比于不同节点基于随机化的神经网络模型参与 AI 学习的实现方式，在上述技术方案中，不同节点可以基于与各自的本地信息对应的神经网络模型参与 AI 学习。从而，在通信系统中的通信节点作为参与 AI 学习的节点的情况下，能够使得通信节点的算力能够应用于学习系统中的 AI 任务，并能够提升不同节点的实现灵活性。

此外，不同节点的本地信息一般是不同的，而在上述技术方案中，第一节点能够基于第一信息确定与本地信息对应的神经网络模型，为此，不同的第一节点均能够基于与本地信息适配的神经网络模型参与 AI 学习，以提升神经网络模型的处理效率。

可选地，不同的第一节点可以为分布式节点，第二节点可以与一个或多个第一节点连接，并作为该一个或多个第一节点的中心节点（或控制节点，管理节点等）。

在第二方面的一种可能的实现方式中，该本地信息包括本地资源开销水平，本地 AI 能力信息、本地的参数规模中的至少一项。

基于上述技术方案，本地信息可以通过多种方式实现，以提供方案实现的灵活性。

应理解，本地资源开销水平可以理解为本地的计算资源的利用率，本地的存储资源的利用率，本地功耗的使用情况等。例如，该本地的计算资源可以包括中央处理器（central processor unit, CPU）、微控制器（micro controller unit, MCU）或图形处理器（graphics processing unit, GPU）等模块/器件的资源；该本地的存储资源可以包括只读存储器（read only memory, ROM）、随机存取存储器（random access memory, RAM）等模块/器件的资源；该本地功耗的使用情况可以包括电池模块的健康度，电池模块当前已使用的电量（或电量的百分比），电池模块当前尚未使用的电量（或电量的百分比）等。

应理解，本地 AI 能力信息可以包括本地与 AI 相关的能力信息，例如与 AI 相关的计算能力、存储能力等。类似的，与 AI 相关的计算能力可以包括 CPU、MCU、GPU 等模块/器件的计算能力，与 AI 相关的存储能力可以包括 ROM、RAM 等模块/器件的存储能力。

在第二方面的一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

基于上述技术方案， N 个配置信息所配置的资源传输的 N 个模型参数包括上述至少一项，以提升方案实现的灵活性。

在第二方面的一种可能的实现方式中， N 为大于 1 的整数，该 N 个模型参数的参数规模不同。

基于上述技术方案，在第一信息指示的 N 个配置信息大于 1 的情况下，该 N 个配置信息所配置的 N 个模型参数的参数规模不同，使得第一节点能够基于本地信息在 N 个配置信息中确定与该本地信息对应的第一配置信息，使得该第一节点能够基于该第一配置信息获得与该本地信息对应的模型参数。

应理解， N 个模型参数的参数规模不同，可以理解为该 N 个模型参数对节点的本地信息的要求不同。例如，以该本地信息包括本地资源开销水平为例， N 个模型参数中参数规模较大的模型参数要求本地资源开销水平低于阈值，以便于提供更多的未被使用的资源处理该参数规模较大的模型参数；又如，以该本地信息包括本地 AI 能力信息为例， N 个模型参数中参数规模较大的模型参数要求本地 AI 能力信息高于阈值，以便于提供更多的 AI 资源处理该参数规模较大的模型参数。

可选地， N 个模型参数的参数规模不同，可以替换为其它的描述，例如 N 个模型参数对应的神经网络层数不同， N 个模型参数对应的神经元不同， N 个模型参数对应的输入输出不同， N 个模型参数的计算复杂度不同，或者， N 个模型参数的计算能力要求不同，或者， N 个模型参数的存储能力要求不同等。

在第二方面的一种可能的实现方式中， N 的取值为 1。

基于上述技术方案，在第一信息指示的 N 个配置信息为一个配置信息（即第一配置信息）的情况下，

该第一配置信息配置的第一模型参数用于配置至少两个第一节点共用的神经网络模型，使得该至少两个第一节点均能够基于该第一配置信息获得共用的神经网络模型，并基于各自的本地信息对该共用的神经网络模型进行知识蒸馏处理，以获得与该本地信息对应的第一神经网络模型。

应理解，至少两个第一节点共用的神经网络模型，可以理解为该至少两个第一节点能够识别/处理的神经网络模型。或者，至少两个第一节点共用的神经网络模型可以表述为：至少两个第一节点通用的神经网络模型，至少两个第一节点能够识别/处理的神经网络模型。

在第二方面的一种可能的实现方式中，该方法还包括：该第二节点发送第二信息，该第二信息指示 M 个配置信息，该 M 个配置信息用于配置 M 个模型参数的传输资源，该 M 个模型参数的参数规模不同， M 为大于 1 的整数；该第二节点基于该 M 个配置信息发送该 M 个模型参数的部分或全部。

基于上述技术方案，第一节点还可以接收指示 M 个配置信息的第二信息，并基于该 M 个配置信息的第二配置信息接收与本地的参数规模相同（或相近）的第二模型参数。换言之，不同的第一节点均可以通过第二信息接收与各自的本地信息的参数规模相同（或相近）的模型参数，并基于所接收的模型参数执行 AI 处理（例如本地训练）。

应理解， M 个配置信息所配置的传输资源用于不同的第一节点接收 M 个模型参数中的其中一个模型参数，使得不同的第一节点后续均可以基于所接收的模型参数执行 AI 处理（例如本地训练）。相应的，该 M 个模型参数可以理解为多个第一节点的全局模型参数，即该 M 个模型参数可以替换为 M 个全局模型参数。

可选地，对于 M 个模型参数中的任一模型参数而言，可以包括梯度和/或权值等。例如，该梯度和/或权值用于初始化的神经网络模型（例如前文描述的第一神经网络模型）的进一步处理，例如神经网络模型的训练，神经网络模型的更新等。

类似地， M 个模型参数的参数规模不同，可以替换为其它的描述，例如 M 个模型参数对应的神经网络层数不同， M 个模型参数对应的神经元不同， M 个模型参数对应的输入输出不同， M 个模型参数的计算复杂度不同，或者， M 个模型参数的计算能力要求不同，或者， M 个模型参数的存储能力要求不同等。

在第二方面的一种可能的实现方式中，该 M 个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的接收功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息。

基于上述技术方案，第二信息指示的 M 个配置信息中的任意一个配置信息可以包括上述至少一项，以提升方案实现的灵活性。

在第二方面的一种可能的实现方式中，该方法还包括：该第二节点发送第三信息，该第三信息指示 K 个配置信息，该 K 个配置信息用于配置 K 个模型参数的传输资源，该 K 个模型参数的参数规模不同， K 为大于 1 的整数；该第二节点基于该 K 个配置信息接收该 K 个模型参数的部分或全部。

基于上述技术方案，第一节点还可以接收指示 K 个配置信息的第三信息，并基于该 K 个配置信息的第三配置信息接收与本地的参数规模相同（或相近）的第三模型参数。换言之，不同的第一节点均可以通过第三信息发送与各自的本地信息的参数规模相同（或相近）的模型参数，使得该模型参数的接收方能够对不同的第一节点发送的模型参数执行 AI 处理（例如参数融合处理）。

应理解， K 个配置信息所配置的传输资源用于不同的第一节点发送 K 个模型参数中的其中一个模型参数，使得该模型参数的接收方能够对不同的第一节点发送的模型参数执行融合操作。相应的，该 K 个模型参数可以理解为多个第一节点的本地模型参数，即该 K 个模型参数可以替换为 K 个本地模型参数。

可选地，对于 K 个模型参数中的任一模型参数而言，可以包括梯度和/或权值等。例如，该梯度和/或权值用于神经网络模型的进一步融合处理，例如深度融合（Deep Fusion），加权平均融合（Weighted Averaging）等。

类似地， K 个模型参数的参数规模不同，可以替换为其它的描述，例如 K 个模型参数对应的神经网络层数不同， K 个模型参数对应的神经元不同， K 个模型参数对应的输入输出不同， K 个模型参数的计算复杂

度不同,或者, K 个模型参数的计算能力要求不同,或者, K 个模型参数的存储能力要求不同等。

在第二方面的一种可能的实现方式中,该 K 个配置信息中的任意一个配置信息包括以下至少一项:

指示该一个配置信息配置的模型参数的参数规模的指示信息;

指示该一个配置信息配置的传输资源的发射功率的指示信息;

5 指示该一个配置信息配置的模型参数的参数类型的指示信息,该参数类型包括梯度和/或权值;

指示该一个配置信息配置的传输资源的频域资源的指示信息;

指示该一个配置信息配置的传输资源的时域资源的指示信息;

指示该一个配置信息的传输资源在该 K 个配置信息配置的传输资源中的索引的指示信息。

基于上述技术方案,第三信息指示的 K 个配置信息中的任意一个配置信息可以包括上述至少一项,

10 以提升方案实现的灵活性。

在第二方面的一种可能的实现方式中,在该第二节点发送第一信息之前,该方法还包括:该第二节点发送第四信息,该第四信息用于指示承载请求信息的资源,该请求信息用于请求该 N 个配置信息;该第二节点在该承载请求信息的资源上接收该请求信息。

基于上述技术方案,第二节点还可以发送用于指示承载请求信息的资源的第四信息,该请求信息用于请求该 N 个配置信息,使得该第一节点能够基于该第四信息指示的资源发送请求信息之后,能够通过第一信息获得该 N 个配置信息。

在第二方面的一种可能的实现方式中,该请求信息包第一正交序列,该第一正交序列对应于该本地信息。

15 可选地,该第一正交序列包含于 P 个正交序列,其中,该 P 个正交序列分别对应于不同的本地信息。换言之,不同第一节点均能够基于各自的本地信息在该 P 个正交序列中确定其中一个正交序列,并向第二节点发送该其中一个正交序列,使得该请求信息的接收方(例如第二节点)能够基于所接收的请求信息获知各个第一节点的本地信息。

20 可选地,在 P 大于1的情况下,不同序列可以是正交的关系,即 P 个序列可以表述为 P 个正交序列,以使该 P 个正交序列通过相同资源传输的情况下也能被接收方所识别。而在其它实现方式中,该 P 个序列也可以是通过其它方式被接收方所识别,例如通过频域资源、时域资源的错开。换言之,上述 P 个正交序列可以替换为 P 个序列,第一正交序列可以替换为第一序列。

25 此外,第二节点作为该请求信息的接收方的情况下,该第二节点能够基于接收的一个或多个正交序列确定各个第一节点的本地信息,可以将各个第一节点的本地信息作为 N 个配置信息以及 M 个配置信息以及 K 个配置信息至少一项的确定依据,使得该第二节点发送的各个配置信息能够与各个第一节点的本地信息相适配。

30 此外,由于不同的第一节点均能够通过该请求信息指示自身的本地参数规模,能够降低开销。

可选地,第一节点除了可以基于第四信息确定承载该第一正交序列的资源之外,该第一节点也可以通过其它方式(例如预配置的方式)确定承载该第一正交序列的资源,此处不做限定。

35 在第二方面的一种可能的实现方式中,在该第二节点发送第一信息之前,该方法还包括:该第二节点发送第五信息,该第五信息用于指示该 N 个模型参数分别对应的本地信息,和/或,该第五信息用于指示该第一信息是否包括该 N 个配置信息。

基于上述技术方案,第二节点还可以发送第五信息,使得第一节点能够通过该第五信息获知该 N 个配置信息的相关信息,并且基于所获知的信息确定是否接收该 N 个配置信息,能够避免第一节点不必要的检测以降低开销。

40 在第二方面的一种可能的实现方式中,在该第二节点发送第三信息之前,该方法还包括:该第二节点发送第六信息,该第六信息用于指示该 M 个模型参数分别对应的本地信息,和/或,该第六信息用于指示该第三信息是否包括该 M 个配置信息。

45 基于上述技术方案,第二节点还可以发送第六信息,以便于通过该第六信息获知该 M 个配置信息的相关信息,并且基于所获知的信息确定是否接收该 M 个配置信息,能够避免第一节点不必要的检测以降低开销。

在第二方面的一种可能的实现方式中,在该第二节点发送第二信息之前,该方法还包括:该第二节点

发送第七信息，该第七信息用于指示该 K 个模型参数分别对应的本地信息，和/或，该第七信息用于指示该第二信息是否包括该 K 个配置信息。

基于上述技术方案，第二节点还可以发送第七信息，以便于通过该第七信息获知该 K 个配置信息的相关信息，并且基于所获知的信息确定是否接收该 K 个配置信息，能够避免第一节点不必要的检测以降低开销。

本申请第三方面提供了一种通信方法，该方法由第一节点执行，或者，该方法由第一节点中的部分组件（例如处理器、芯片或芯片系统等）执行，或者该方法还可以由能实现全部或部分第一节点功能的逻辑模块或软件实现。在第三方面及其可能的实现方式中，以该方法由第一节点执行为例进行描述。在该方法中，第一节点接收第二信息，该第二信息指示 M 个配置信息，该 M 个配置信息用于配置 M 个模型参数的传输资源，该 M 个模型参数的参数规模不同， M 为大于 1 的整数；该第一节点基于该 M 个配置信息中的第二配置信息接收第二模型参数，其中，该第二模型参数的参数规模与本地的参数规模相同（或相近，例如该第二模型参数的参数规模与本地的参数规模的差低于阈值）。

基于上述技术方案，第一节点接收的第二信息指示 M 个配置信息，该 M 个配置信息用于配置 M 个模型参数的资源，此后，该第一节点基于该 M 个配置信息中的第二配置信息所配置的资源接收第二模型参数，并且，该第二模型参数与本地的参数规模相同（或相近）。换言之，不同的第一节点均可以基于该第一信息获得与本地的参数规模相同（或相近）的模型参数，后续均可以基于与本地的参数规模相同（或相近）的模型参数参与 AI 学习。相比于不同节点基于随机化的模型参数参与 AI 学习的实现方式，在上述技术方案中，不同节点可以基于与本地的参数规模相同（或相近）的模型参数参与 AI 学习。从而，在通信系统中的通信节点作为参与 AI 学习的节点的情况下，能够使得通信节点的算力能够应用于学习系统中的 AI 任务，并能够提升不同节点的实现灵活性。

此外，不同节点的本地参数规模可能是不同的，在上述技术方案中，第一节点能够基于与本地的参数规模相同（或相近）的模型参数参与 AI 学习，为此，不同的第一节点均能够基于与本地信息适配的模型参数参与 AI 学习，以提升神经网络模型的处理效率。

在第三方面的一种可能的实现方式中，第一节点可以基于该本地信息在该 M 个配置信息中确定该第二配置信息，该本地信息包括本地资源开销水平，本地 AI 能力信息、本地的参数规模中的至少一项。

基于上述技术方案，本地信息可以通过多种方式实现，以提供方案实现的灵活性。

应理解，本地资源开销水平可以理解为本地的计算资源的利用率，本地的存储资源的利用率，本地功耗的使用情况等。例如，该本地的计算资源可以包括中央处理器（central processor unit, CPU）、微控制器（micro controller unit, MCU）或图形处理器（graphics processing unit, GPU）等模块/器件的资源；该本地的存储资源可以包括只读存储器（read only memory, ROM）、随机存取存储器（random access memory, RAM）等模块/器件的资源；该本地功耗的使用情况可以包括电池模块的健康度，电池模块当前已使用的电量（或电量的百分比），电池模块当前尚未使用的电量（或电量的百分比）等。

应理解，本地 AI 能力信息可以包括本地与 AI 相关的能力信息，例如与 AI 相关的计算能力、存储能力等。类似的，与 AI 相关的计算能力可以包括 CPU、MCU、GPU 等模块/器件的计算能力，与 AI 相关的存储能力可以包括 ROM、RAM 等模块/器件的存储能力。

应理解， M 个配置信息所配置的传输资源用于不同的第一节点接收 M 个模型参数中的其中一个模型参数，使得不同的第一节点后续均可以基于所接收的模型参数执行 AI 处理（例如本地训练）。相应的，该 M 个模型参数可以理解为多个第一节点的全局模型参数，即该 M 个模型参数可以替换为 M 个全局模型参数。

可选地，对于 M 个模型参数中的任一模型参数而言，可以包括梯度和/或权值等。例如，该梯度和/或权值用于初始化的神经网络模型（例如前文描述的第一神经网络模型）的进一步处理，例如神经网络模型的训练，神经网络模型的更新等。

类似地， M 个模型参数的参数规模不同，可以替换为其它的描述，例如 M 个模型参数对应的神经网络层数不同， M 个模型参数对应的神经元不同， M 个模型参数对应的输入输出不同， M 个模型参数的计算复杂度不同，或者， M 个模型参数的计算能力要求不同，或者， M 个模型参数的存储能力要求不同等。

在第三方面的一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

基于上述技术方案， M 个配置信息所配置的资源传输的 M 个模型参数包括上述至少一项，以提升方案实现的灵活性。

- 5 在第三方面的一种可能的实现方式中，该 M 个配置信息中的任意一个配置信息包括以下至少一项：
指示该一个配置信息配置的模型参数的参数规模的指示信息；
指示该一个配置信息配置的传输资源的接收功率的指示信息；
指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；
指示该一个配置信息配置的传输资源的频域资源的指示信息；
10 指示该一个配置信息配置的传输资源的时域资源的指示信息。

基于上述技术方案，第一节点能够基于本地信息在 M 个配置信息中确定与该本地信息对应的第二配置信息，使得该第一节点能够基于该第二配置信息获得与该本地信息对应的模型参数。

- 在第三方面的一种可能的实现方式中，该方法还包括：该第一节点接收第三信息，该第三信息指示 K 个配置信息，该 K 个配置信息用于配置 K 个模型参数的传输资源，该 K 个模型参数的参数规模不同， K 为
15 大于 1 的整数；该第一节点基于该 K 个配置信息中的第三配置信息发送第三模型参数，其中，该第三模型参数的参数规模与本地的参数规模相同（或相近，例如该第三模型参数的参数规模与本地的参数规模的差低于阈值）。

- 基于上述技术方案，第一节点还可以接收指示 K 个配置信息的第三信息，并基于该 K 个配置信息的第三配置信息接收与本地的参数规模相同（或相近）的第三模型参数。换言之，不同的第一节点均可以通过
20 第三信息发送与各自的本地信息的参数规模相同（或相近）的模型参数，使得该模型参数的接收方能够对不同的第一节点发送的模型参数执行 AI 处理（例如参数融合处理）。

应理解， K 个配置信息所配置的传输资源用于不同的第一节点发送 K 个模型参数中的其中一个模型参数，使得该模型参数的接收方能够对不同的第一节点发送的模型参数执行融合操作。相应的，该 K 个模型参数可以理解为多个第一节点的本地模型参数，即该 K 个模型参数可以替换为 K 个本地模型参数。

- 25 可选地，对于 K 个模型参数中的任一模型参数而言，可以包括梯度和/或权值等。例如，该梯度和/或权值用于神经网络模型的进一步融合处理，例如深度融合（Deep Fusion），加权平均融合（Weighted Averaging）等。

- 类似地， K 个模型参数的参数规模不同，可以替换为其它的描述，例如 K 个模型参数对应的神经网络
30 层数不同， K 个模型参数对应的神经元不同， K 个模型参数对应的输入输出不同， K 个模型参数的计算复杂度不同，或者， K 个模型参数的计算能力要求不同，或者， K 个模型参数的存储能力要求不同等。

- 在第三方面的一种可能的实现方式中，该 K 个配置信息中的任意一个配置信息包括以下至少一项：
指示该一个配置信息配置的模型参数的参数规模的指示信息；
指示该一个配置信息配置的传输资源的发射功率的指示信息；
指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；
35 指示该一个配置信息配置的传输资源的频域资源的指示信息；
指示该一个配置信息配置的传输资源的时域资源的指示信息；
指示该一个配置信息的传输资源在该 K 个配置信息配置的传输资源中的索引的指示信息。

基于上述技术方案，第三信息指示的 K 个配置信息中的任意一个配置信息可以包括上述至少一项，以提升方案实现的灵活性。

- 40 在第三方面的一种可能的实现方式中，在该第一节点接收第三信息之前，该方法还包括：该第一节点接收第六信息，该第六信息用于指示该 M 个模型参数分别对应的本地信息，和/或，该第六信息用于指示该第三信息是否包括该 M 个配置信息。

- 基于上述技术方案，第一节点还可以接收第六信息，以便于通过该第六信息获知该 M 个配置信息的相关信息，并且基于所获知的信息确定是否接收该 M 个配置信息，能够避免第一节点不必要的检测以降低开销。
45

在第三方面的一种可能的实现方式中，在该第一节点接收第二信息之前，该方法还包括：该第一节点

接收第七信息，该第七信息用于指示该 K 个模型参数分别对应的本地信息，和/或，该第七信息用于指示该第二信息是否包括该 K 个配置信息。

基于上述技术方案，第一节点还可以接收第七信息，以便于通过该第七信息获知该 K 个配置信息的相关信息，并且基于所获知的信息确定是否接收该 K 个配置信息，能够避免第一节点不必要的检测以降低开销。

本申请第四方面提供了一种通信方法，该方法由第二节点执行，或者，该方法由第二节点中的部分组件（例如处理器、芯片或芯片系统等）执行，或者该方法还可以由能实现全部或部分第二节点功能的逻辑模块或软件实现。在第四方面及其可能的实现方式中，以该方法由第二节点执行为例进行描述。在该方法中，第二节点确定第二信息，该第二信息指示 M 个配置信息，该 M 个配置信息用于配置 M 个模型参数的传输资源，该 M 个模型参数的参数规模不同， M 为大于 1 的整数；该第二节点发送第二信息。

基于上述技术方案，第二节点发送的第二信息指示 M 个配置信息，该 M 个配置信息用于配置 M 个模型参数的资源，此后，第一节点接收第二信息之后，该第一节点基于该 M 个配置信息中的第二配置信息所配置的资源接收第二模型参数，并且，该第二模型参数与本地的参数规模相同（或相近）。换言之，不同的第一节点均可以基于该第一信息获得与本地的参数规模相同（或相近）的模型参数，后续均可以基于与本地的参数规模相同（或相近）的模型参数参与 AI 学习。相比于不同节点基于随机化的模型参数参与 AI 学习的实现方式，在上述技术方案中，不同节点可以基于与本地的参数规模相同（或相近）的模型参数参与 AI 学习。从而，在通信系统中的通信节点作为参与 AI 学习的节点的情况下，能够使得通信节点的算力能够应用于学习系统中的 AI 任务，并能够提升不同节点的实现灵活性。

此外，不同节点的本地参数规模可能是不同的，在上述技术方案中，第一节点能够基于与本地的参数规模相同（或相近）的模型参数参与 AI 学习，为此，不同的第一节点均能够基于与本地信息适配的模型参数参与 AI 学习，以提升神经网络模型的处理效率。

应理解， M 个配置信息所配置的传输资源用于不同的第一节点接收 M 个模型参数中的其中一个模型参数，使得不同的第一节点后续均可以基于所接收的模型参数执行 AI 处理（例如本地训练）。相应的，该 M 个模型参数可以理解为多个第一节点的全局模型参数，即该 M 个模型参数可以替换为 M 个全局模型参数。

可选地，对于 M 个模型参数中的任一模型参数而言，可以包括梯度和/或权值等。例如，该梯度和/或权值用于初始化的神经网络模型（例如前文描述的第一神经网络模型）的进一步处理，例如神经网络模型的训练，神经网络模型的更新等。

类似地， M 个模型参数的参数规模不同，可以替换为其它的描述，例如 M 个模型参数对应的神经网络层数不同， M 个模型参数对应的神经元不同， M 个模型参数对应的输入输出不同， M 个模型参数的计算复杂度不同，或者， M 个模型参数的计算能力要求不同，或者， M 个模型参数的存储能力要求不同等。

在第四方面的一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

基于上述技术方案， M 个配置信息所配置的资源传输的 M 个模型参数包括上述至少一项，以提升方案实现的灵活性。

在第四方面的一种可能的实现方式中，该 M 个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的接收功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息。

基于上述技术方案，第一节点能够基于本地信息在 M 个配置信息中确定与该本地信息对应的第二配置信息，使得该第一节点能够基于该第二配置信息获得与该本地信息对应的模型参数。

在第四方面的一种可能的实现方式中，该方法还包括：该第二节点发送第三信息，该第三信息指示 K 个配置信息，该 K 个配置信息用于配置 K 个模型参数的传输资源，该 K 个模型参数的参数规模不同， K 为

大于 1 的整数。

基于上述技术方案，第二节点还可以发送指示 K 个配置信息的第三信息，使得第一节点能够基于该 K 个配置信息的第三配置信息接收与本地的参数规模相同（或相近）的第三模型参数。换言之，不同的第一节点均可以通过第三信息发送与各自的本地信息的参数规模相同（或相近）的模型参数，使得该模型参数的接收方能够对不同的第一节点发送的模型参数执行 AI 处理（例如参数融合处理）。

应理解，K 个配置信息所配置的传输资源用于不同的第一节点发送 K 个模型参数中的其中一个模型参数，使得该模型参数的接收方能够对不同的第一节点发送的模型参数执行融合操作。相应的，该 K 个模型参数可以理解多个第一节点的本地模型参数，即该 K 个模型参数可以替换为 K 个本地模型参数。

可选地，对于 K 个模型参数中的任一模型参数而言，可以包括梯度和/或权值等。例如，该梯度和/或权值用于神经网络模型的进一步融合处理，例如深度融合（Deep Fusion），加权平均融合（Weighted Averaging）等。

类似地，K 个模型参数的参数规模不同，可以替换为其它的描述，例如 K 个模型参数对应的神经网络层数不同，K 个模型参数对应的神经元不同，K 个模型参数对应的输入输出不同，K 个模型参数的计算复杂度不同，或者，K 个模型参数的计算能力要求不同，或者，K 个模型参数的存储能力要求不同等。

在第四方面的一种可能的实现方式中，该 K 个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的发射功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息；

指示该一个配置信息的传输资源在该 K 个配置信息配置的传输资源中的索引的指示信息。

基于上述技术方案，第三信息指示的 K 个配置信息中的任意一个配置信息可以包括上述至少一项，以提升方案实现的灵活性。

在第四方面的一种可能的实现方式中，在该第二节点发送第三信息之前，该方法还包括：该第二节点发送第六信息，该第六信息用于指示该 M 个模型参数分别对应的本地信息，和/或，该第六信息用于指示该第三信息是否包括该 M 个配置信息。

基于上述技术方案，第二节点还可以发送第六信息，以便于通过该第六信息获知该 M 个配置信息的相关信息，并且基于所获知的信息确定是否接收该 M 个配置信息，能够避免第一节点不必要的检测以降低开销。

在第四方面的一种可能的实现方式中，在该第二节点发送第二信息之前，该方法还包括：该第二节点发送第七信息，该第七信息用于指示该 K 个模型参数分别对应的本地信息，和/或，该第七信息用于指示该第二信息是否包括该 K 个配置信息。

基于上述技术方案，第二节点还可以发送第七信息，以便于通过该第七信息获知该 K 个配置信息的相关信息，并且基于所获知的信息确定是否接收该 K 个配置信息，能够避免第一节点不必要的检测以降低开销。

本申请第五方面提供了一种通信方法，该方法由第一节点执行，或者，该方法由第一节点中的部分组件（例如处理器、芯片或芯片系统等）执行，或者该方法还可以由能实现全部或部分第一节点功能的逻辑模块或软件实现。在第五方面及其可能的实现方式中，以该方法由第一节点执行为例进行描述。在该方法中，第一节点接收第三信息，该第三信息指示 K 个配置信息，该 K 个配置信息用于配置 K 个模型参数的传输资源，该 K 个模型参数的参数规模不同，K 为大于 1 的整数；该第一节点基于该 K 个配置信息中的第三配置信息发送第三模型参数，其中，该第三模型参数的参数规模与本地的参数规模相同（或相近，例如该第三模型参数的参数规模与本地的参数规模的差低于阈值）。

基于上述技术方案，第一节点接收的第三信息指示 K 个配置信息，该 K 个配置信息用于配置 K 个模型参数的资源，此后，该第一节点基于该 K 个配置信息中的第三配置信息所配置的资源发送第三模型参数，并且，该第三模型参数与本地的参数规模相同（或相近）。换言之，不同的第一节点均可以基于该第一信

息获得与本地的参数规模相同（或相近）的模型参数，后续均可以基于与本地的参数规模相同（或相近）的模型参数参与 AI 学习。相比于不同节点基于随机化的模型参数参与 AI 学习的实现方式，在上述技术方案中，不同节点可以基于与本地的参数规模相同（或相近）的模型参数参与 AI 学习。从而，在通信系统中的通信节点作为参与 AI 学习的节点的情况下，能够使得通信节点的算力能够应用于学习系统中的 AI 任务，并能够提升不同节点的实现灵活性。

此外，不同节点的本地参数规模可能是不同的，在上述技术方案中，第一节点能够基于与本地的参数规模相同（或相近）的模型参数参与 AI 学习，为此，不同的第一节点均能够基于与本地信息适配的模型参数参与 AI 学习，以提升神经网络模型的处理效率。

在第五方面的一种可能的实现方式中，第一节点可以基于该本地信息在该 M 个配置信息中确定该第二配置信息。

基于上述技术方案，本地信息可以通过多种方式实现，以提供方案实现的灵活性。

应理解，本地资源开销水平可以理解为本地的计算资源的利用率，本地的存储资源的利用率，本地功耗的使用情况等。例如，该本地的计算资源可以包括中央处理器（central processor unit, CPU）、微控制器（micro controller unit, MCU）或图形处理器（graphics processing unit, GPU）等模块/器件的资源；该本地的存储资源可以包括只读存储器（read only memory, ROM）、随机存取存储器（random access memory, RAM）等模块/器件的资源；该本地功耗的使用情况可以包括电池模块的健康度，电池模块当前已使用的电量（或电量的百分比），电池模块当前尚未使用的电量（或电量的百分比）等。

应理解，本地 AI 能力信息可以包括本地与 AI 相关的能力信息，例如与 AI 相关的计算能力、存储能力等。类似的，与 AI 相关的计算能力可以包括 CPU、MCU、GPU 等模块/器件的计算能力，与 AI 相关的存储能力可以包括 ROM、RAM 等模块/器件的存储能力。

应理解，K 个配置信息所配置的传输资源用于不同的第一节点发送 K 个模型参数中的其中一个模型参数，使得该模型参数的接收方能够对不同的第一节点发送的模型参数执行融合操作。相应的，该 K 个模型参数可以理解为多个第一节点的本地模型参数，即该 K 个模型参数可以替换为 K 个本地模型参数。

可选地，对于 K 个模型参数中的任一模型参数而言，可以包括梯度和/或权值等。例如，该梯度和/或权值用于神经网络模型的进一步融合处理，例如深度融合（Deep Fusion），加权平均融合（Weighted Averaging）等。

类似地，K 个模型参数的参数规模不同，可以替换为其它的描述，例如 K 个模型参数对应的神经网络层数不同，K 个模型参数对应的神经元不同，K 个模型参数对应的输入输出不同，K 个模型参数的计算复杂度不同，或者，K 个模型参数的计算能力要求不同，或者，K 个模型参数的存储能力要求不同等。

在第五方面的一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

基于上述技术方案，K 个配置信息所配置的资源传输的 K 个模型参数包括上述至少一项，以提升方案实现的灵活性。

在第五方面的一种可能的实现方式中，该 K 个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的发射功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息；

指示该一个配置信息的传输资源在该 K 个配置信息配置的传输资源中的索引的指示信息。

基于上述技术方案，第一节点能够基于本地信息在 K 个配置信息中确定与该本地信息对应的第三配置信息，使得该第一节点能够基于该第三配置信息发送与该本地信息对应的模型参数。

在第五方面的一种可能的实现方式中，在该第一节点接收第二信息之前，该方法还包括：该第一节点接收第七信息，该第七信息用于指示该 K 个模型参数分别对应的本地信息，和/或，该第七信息用于指示该第二信息是否包括该 K 个配置信息。

基于上述技术方案，第一节点还可以接收第七信息，以便于通过该第七信息获知该 K 个配置信息的相

关信息，并且基于所获知的信息确定是否接收该 K 个配置信息，能够避免第一节点不必要的检测以降低开销。

本申请第六方面提供了一种通信方法，该方法由第二节点执行，或者，该方法由第二节点中的部分组件（例如处理器、芯片或芯片系统等）执行，或者该方法还可以由能实现全部或部分第二节点功能的逻辑模块或软件实现。在第六方面及其可能的实现方式中，以该方法由第二节点执行为例进行描述。在该方法中，第二节点确定第三信息，该第三信息指示 K 个配置信息，该 K 个配置信息用于配置 K 个模型参数的传输资源，该 K 个模型参数的参数规模不同，K 为大于 1 的整数；该第二节点发送第三信息。

基于上述技术方案，第二节点发送的第三信息指示 K 个配置信息，该 K 个配置信息用于配置 K 个模型参数的资源，此后，第一节点接收第三信息之后，该第一节点基于该 K 个配置信息中的第二配置信息所配置的资源发送第三模型参数，并且，该第三模型参数与本地的参数规模相同（或相近）。换言之，不同的第一节点均可以基于该第一信息获得与本地的参数规模相同（或相近）的模型参数，后续均可以基于与本地的参数规模相同（或相近）的模型参数参与 AI 学习。相比于不同节点基于随机化的模型参数参与 AI 学习的实现方式，在上述技术方案中，不同节点可以基于与本地的参数规模相同（或相近）的模型参数参与 AI 学习。从而，在通信系统中的通信节点作为参与 AI 学习的节点的情况下，能够使得通信节点的算力能够应用于学习系统中的 AI 任务，并能够提升不同节点的实现灵活性。

此外，不同节点的本地参数规模可能是不同的，在上述技术方案中，第一节点能够基于与本地的参数规模相同（或相近）的模型参数参与 AI 学习，为此，不同的第一节点均能够基于与本地信息适配的模型参数参与 AI 学习，以提升神经网络模型的处理效率。

应理解，K 个配置信息所配置的传输资源用于不同的第一节点发送 K 个模型参数中的其中一个模型参数，使得该模型参数的接收方能够对不同的第一节点发送的模型参数执行融合操作。相应的，该 K 个模型参数可以理解为多个第一节点的本地模型参数，即该 K 个模型参数可以替换为 K 个本地模型参数。

可选地，对于 K 个模型参数中的任一模型参数而言，可以包括梯度和/或权值等。例如，该梯度和/或权值用于神经网络模型的进一步融合处理，例如深度融合（Deep Fusion），加权平均融合（Weighted Averaging）等。

类似地，K 个模型参数的参数规模不同，可以替换为其它的描述，例如 K 个模型参数对应的神经网络层数不同，K 个模型参数对应的神经元不同，K 个模型参数对应的输入输出不同，K 个模型参数的计算复杂度不同，或者，K 个模型参数的计算能力要求不同，或者，K 个模型参数的存储能力要求不同等。

在第六方面的一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

基于上述技术方案，K 个配置信息所配置的资源传输的 K 个模型参数包括上述至少一项，以提升方案实现的灵活性。

在第六方面的一种可能的实现方式中，该 K 个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的发射功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息；

指示该一个配置信息的传输资源在该 K 个配置信息配置的传输资源中的索引的指示信息。

基于上述技术方案，第一节点能够基于本地信息在 K 个配置信息中确定与该本地信息对应的第三配置信息，使得该第一节点能够基于该第三配置信息发送与该本地信息对应的模型参数。

在第六方面的一种可能的实现方式中，在该第二节点发送第二信息之前，该方法还包括：该第二节点发送第七信息，该第七信息用于指示该 K 个模型参数分别对应的本地信息，和/或，该第七信息用于指示该第二信息是否包括该 K 个配置信息。

基于上述技术方案，第二节点还可以发送第七信息，以便于通过该第七信息获知该 K 个配置信息的相关信息，并且基于所获知的信息确定是否接收该 K 个配置信息，能够避免第一节点不必要的检测以降低开

销。

本申请第七方面提供了一种通信装置，该装置为第一节点，或者，该装置为第一节点中的部分组件（例如处理器、芯片或芯片系统等），或者，该装置还可以为能够实现全部或部分第一节点功能的逻辑模块或软件。在第七方面及其可能的实现方式中，以该通信装置为第一节点执行为例进行描述，该第一节点可以为终端设备或网络设备。

该装置包括处理单元和收发单元；该收发单元用于接收第一信息，该第一信息指示 N 个配置信息，该 N 个配置信息用于配置 N 个模型参数的资源， N 为正整数；该处理单元用于确定该 N 个配置信息中的第一配置信息；该收发单元还用于基于该第一配置信息接收第一模型参数，该第一模型参数用于确定与本地信息对应的第一神经网络模型。

在第七方面的一种可能的实现方式中，该本地信息包括本地资源开销水平，本地 AI 能力信息、本地的参数规模中的至少一项。

在第七方面的一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

在第七方面的一种可能的实现方式中， N 为大于 1 的整数，该 N 个模型参数的参数规模不同，该处理单元还用于基于该本地信息在该 N 个配置信息中确定该第一配置信息。

在第七方面的一种可能的实现方式中， N 的取值为 1，其中，该第一模型参数用于配置至少两个第一节点共用的神经网络模型；该处理单元还用于基于该本地信息将该至少两个第一节点共用的神经网络模型进行知识蒸馏处理，得到该第一神经网络模型。

在第七方面的一种可能的实现方式中，该收发单元还用于接收第二信息，该第二信息指示 M 个配置信息，该 M 个配置信息用于配置 M 个模型参数的传输资源，该 M 个模型参数的参数规模不同， M 为大于 1 的整数；该处理单元还用于确定该 M 个配置信息中的第二配置信息；该收发单元还用于基于该第二配置信息接收第二模型参数，其中，该第二模型参数的参数规模与本地的参数规模相同（或相近）。

在第七方面的一种可能的实现方式中，该 M 个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的接收功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息。

在第七方面的一种可能的实现方式中，该收发单元还用于接收第三信息，该第三信息指示 K 个配置信息，该 K 个配置信息用于配置 K 个模型参数的传输资源，该 K 个模型参数的参数规模不同， K 为大于 1 的整数；该处理单元还用于确定该 K 个配置信息中的第三配置信息；该收发单元还用于基于该第三配置信息发送第三模型参数，其中，该第三模型参数的参数规模与本地的参数规模相同（或相近）。

在第七方面的一种可能的实现方式中，该 K 个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的发射功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息；

指示该一个配置信息的传输资源在该 K 个配置信息配置的传输资源中的索引的指示信息。

在第七方面的一种可能的实现方式中，该收发单元还用于接收第四信息，该第四信息用于指示承载请求信息的资源，该请求信息用于请求该 N 个配置信息；该收发单元还用于在该承载请求信息的资源上发送该请求信息。

在第七方面的一种可能的实现方式中，该请求信息包括第一正交序列，该第一正交序列对应于该本地信息。

在第七方面的一种可能的实现方式中，该收发单元还用于接收第五信息，该第五信息用于指示该 N 个

模型参数分别对应的本地信息, 和/或, 该第五信息用于指示该第一信息是否包括该 N 个配置信息。

本申请第七方面中, 通信装置的组成模块还可以用于执行第一方面的各个可能实现方式中所执行的步骤, 并实现相应的技术效果, 具体均可以参阅第一方面, 此处不再赘述。

5 本申请第八方面提供了一种通信装置, 该装置为第二节点, 或者, 该装置为第二节点中的部分组件 (例如处理器、芯片或芯片系统等), 或者, 该装置还可以为能够实现全部或部分第二节点功能的逻辑模块或软件。在第八方面及其可能的实现方式中, 以该通信装置为第二节点执行为例进行描述, 该第二节点可以为终端设备或网络设备。

该装置包括处理单元和收发单元; 该处理单元用于确定第一信息, 该第一信息指示 N 个配置信息, 该
10 N 个配置信息用于配置 N 个模型参数的资源, N 为正整数; 其中, 该 N 个模型参数中的一个模型参数用于确定与第一节点的本地信息对应的神经网络模型; 该收发单元用于发送该第一信息。

在第八方面的一种可能的实现方式中, 该本地信息包括本地资源开销水平, 本地 AI 能力信息、本地的参数规模中的至少一项。

15 在第八方面的一种可能的实现方式中, 该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

在第八方面的一种可能的实现方式中, N 为大于 1 的整数, 该 N 个模型参数的参数规模不同。

在第八方面的一种可能的实现方式中, N 的取值为 1。

在第八方面的一种可能的实现方式中, 该收发单元还用于发送第二信息, 该第二信息指示 M 个配置信息, 该 M 个配置信息用于配置 M 个模型参数的传输资源, 该 M 个模型参数的参数规模不同, M 为大于 1
20 的整数; 该收发单元还用于基于该 M 个配置信息发送该 M 个模型参数的部分或全部。

在第八方面的一种可能的实现方式中, 该 M 个配置信息中的任意一个配置信息包括以下至少一项:

指示该一个配置信息配置的模型参数的参数规模的指示信息;

指示该一个配置信息配置的传输资源的接收功率的指示信息;

指示该一个配置信息配置的模型参数的参数类型的指示信息, 该参数类型包括梯度和/或权值;

25 指示该一个配置信息配置的传输资源的频域资源的指示信息;

指示该一个配置信息配置的传输资源的时域资源的指示信息。

在第八方面的一种可能的实现方式中, 该收发单元还用于发送第三信息, 该第三信息指示 K 个配置信息, 该 K 个配置信息用于配置 K 个模型参数的传输资源, 该 K 个模型参数的参数规模不同, K 为大于 1 的
30 整数; 该收发单元还用于基于该 K 个配置信息接收该 K 个模型参数的部分或全部。

在第八方面的一种可能的实现方式中, 该 K 个配置信息中的任意一个配置信息包括以下至少一项:

指示该一个配置信息配置的模型参数的参数规模的指示信息;

指示该一个配置信息配置的传输资源的发射功率的指示信息;

指示该一个配置信息配置的模型参数的参数类型的指示信息, 该参数类型包括梯度和/或权值;

35 指示该一个配置信息配置的传输资源的频域资源的指示信息;

指示该一个配置信息配置的传输资源的时域资源的指示信息;

指示该一个配置信息的传输资源在该 K 个配置信息配置的传输资源中的索引的指示信息。

在第八方面的一种可能的实现方式中, 该收发单元还用于发送第四信息, 该第四信息用于指示承载请求信息的资源, 该请求信息用于请求该 N 个配置信息; 该收发单元还用于在该承载请求信息的资源上接收
40 该请求信息。

在第八方面的一种可能的实现方式中, 该请求信息包括第一正交序列, 该第一正交序列对应于该本地信息。

在第八方面的一种可能的实现方式中, 该收发单元还用于发送第五信息, 该第五信息用于指示该 N 个模型参数分别对应的本地信息, 和/或, 该第五信息用于指示该第一信息是否包括该 N 个配置信息。

45 本申请第八方面中, 通信装置的组成模块还可以用于执行第二方面的各个可能实现方式中所执行的步骤, 并实现相应的技术效果, 具体均可以参阅第二方面, 此处不再赘述。

本申请第九方面提供了一种通信装置，该装置为第一节点，或者，该装置为第一节点中的部分组件（例如处理器、芯片或芯片系统等），或者，该装置还可以为能够实现全部或部分第一节点功能的逻辑模块或软件。在第九方面及其可能的实现方式中，以该通信装置为第一节点执行为例进行描述，该第一节点可以为终端设备或网络设备。

5 该装置包括处理单元和收发单元；该收发单元用于接收第二信息，该第二信息指示 M 个配置信息，该 M 个配置信息用于配置 M 个模型参数的传输资源，该 M 个模型参数的参数规模不同， M 为大于 1 的整数；该处理单元用于确定 M 个配置信息中的第二配置信息；该收发单元还用于基于该第二配置信息接收第二模型参数，其中，该第二模型参数的参数规模与本地的参数规模相同（或相近）。

10 在第九方面的一种可能的实现方式中，该本地信息包括本地资源开销水平，本地 AI 能力信息、本地的参数规模中的至少一项。

在第九方面的一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

在第九方面的一种可能的实现方式中，该 M 个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

15 指示该一个配置信息配置的传输资源的接收功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息。

20 在第九方面的一种可能的实现方式中，该收发单元还用于接收第三信息，该第三信息指示 K 个配置信息，该 K 个配置信息用于配置 K 个模型参数的传输资源，该 K 个模型参数的参数规模不同， K 为大于 1 的整数；该收发单元还用于基于该 K 个配置信息中的第三配置信息发送第三模型参数，其中，该第三模型参数的参数规模与本地的参数规模相同（或相近）。

在第九方面的一种可能的实现方式中，该 K 个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

25 指示该一个配置信息配置的传输资源的发射功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息；

指示该一个配置信息的传输资源在该 K 个配置信息配置的传输资源中的索引的指示信息。

30 本申请第九方面中，通信装置的组成模块还可以用于执行第三方面的各个可能实现方式中所执行的步骤，并实现相应的技术效果，具体均可以参阅第三方面，此处不再赘述。

35 本申请第十方面提供了一种通信装置，该装置为第二节点，或者，该装置为第二节点中的部分组件（例如处理器、芯片或芯片系统等），或者，该装置还可以为能够实现全部或部分第二节点功能的逻辑模块或软件。在第十方面及其可能的实现方式中，以该通信装置为第二节点执行为例进行描述，该第二节点可以为终端设备或网络设备。

该装置包括处理单元和收发单元；该处理单元用于确定第二信息，该第二信息指示 M 个配置信息，该 M 个配置信息用于配置 M 个模型参数的传输资源，该 M 个模型参数的参数规模不同， M 为大于 1 的整数；该收发单元用于发送第二信息。

40 在第十方面的一种可能的实现方式中，该本地信息包括本地资源开销水平，本地 AI 能力信息、本地的参数规模中的至少一项。

在第十方面的一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

在第十方面的一种可能的实现方式中，该 M 个配置信息中的任意一个配置信息包括以下至少一项：

45 指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的接收功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息。

在第十方面的一种可能的实现方式中，该处理单元还用于确定第三信息，该第三信息指示 K 个配置信息，该 K 个配置信息用于配置 K 个模型参数的传输资源，该 K 个模型参数的参数规模不同， K 为大于 1 的整数；

在第十方面的一种可能的实现方式中，该 K 个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的发射功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息；

指示该一个配置信息的传输资源在该 K 个配置信息配置的传输资源中的索引的指示信息。

本申请第十方面中，通信装置的组成模块还可以用于执行第四方面的各个可能实现方式中所执行的步骤，并实现相应的技术效果，具体均可以参阅第四方面，此处不再赘述。

本申请第十一方面提供了一种通信装置，该装置为第一节点，或者，该装置为第一节点中的部分组件（例如处理器、芯片或芯片系统等），或者，该装置还可以为能够实现全部或部分第一节点功能的逻辑模块或软件。在第十一方面及其可能的实现方式中，以该通信装置为第一节点执行为例进行描述，该第一节点可以为终端设备或网络设备。

该装置包括处理单元和收发单元；该收发单元用于接收第三信息，该第三信息指示 K 个配置信息，该 K 个配置信息用于配置 K 个模型参数的传输资源，该 K 个模型参数的参数规模不同， K 为大于 1 的整数；该处理单元用于确定该 K 个配置信息中的第三配置信息；该收发单元还用于基于该第三配置信息发送第三模型参数，其中，该第三模型参数的参数规模与本地的参数规模相同（或相近）。

在第十一方面的一种可能的实现方式中，该本地信息包括本地资源开销水平，本地 AI 能力信息、本地的参数规模中的至少一项。

在第十一方面的一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

在第十一方面的一种可能的实现方式中，该 K 个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的发射功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息；

指示该一个配置信息的传输资源在该 K 个配置信息配置的传输资源中的索引的指示信息。

本申请第十一方面中，通信装置的组成模块还可以用于执行第五方面的各个可能实现方式中所执行的步骤，并实现相应的技术效果，具体均可以参阅第五方面，此处不再赘述。

本申请第十二方面提供了一种通信装置，该装置为第二节点，或者，该装置为第二节点中的部分组件（例如处理器、芯片或芯片系统等），或者，该装置还可以为能够实现全部或部分第二节点功能的逻辑模块或软件。在第十二方面及其可能的实现方式中，以该通信装置为第二节点执行为例进行描述，该第二节点可以为终端设备或网络设备。

该装置包括处理单元和收发单元；该处理单元用于确定第三信息，该第三信息指示 K 个配置信息，该 K 个配置信息用于配置 K 个模型参数的传输资源，该 K 个模型参数的参数规模不同， K 为大于 1 的整数；该收发单元用于发送第三信息。

在第十二方面的一种可能的实现方式中，该本地信息包括本地资源开销水平，本地 AI 能力信息、本地

的参数规模中的至少一项。

在第十二方面的一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

在第十二方面的一种可能的实现方式中，该 K 个配置信息中的任意一个配置信息包括以下至少一项：

- 5 指示该一个配置信息配置的模型参数的参数规模的指示信息；
指示该一个配置信息配置的传输资源的发射功率的指示信息；
指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；
指示该一个配置信息配置的传输资源的频域资源的指示信息；
指示该一个配置信息配置的传输资源的时域资源的指示信息；
10 指示该一个配置信息的传输资源在该 K 个配置信息配置的传输资源中的索引的指示信息。

本申请第十二方面中，通信装置的组成模块还可以用于执行第六方面的各个可能实现方式中所执行的步骤，并实现相应的技术效果，具体均可以参阅第六方面，此处不再赘述。

本申请实施例第十三方面提供了一种通信装置，包括至少一个处理器，所述至少一个处理器与存储器耦合；该存储器用于存储程序或指令；该至少一个处理器用于执行该程序或指令，以使该装置实现前述第一方面或第一方面任意一种可能的实现方式所述的方法。

本申请实施例第十四方面提供了一种通信装置，包括至少一个处理器，所述至少一个处理器与存储器耦合；该存储器用于存储程序或指令；该至少一个处理器用于执行该程序或指令，以使该装置实现前述第二方面或第二方面任意一种可能的实现方式所述的方法。

本申请实施例第十五方面提供了一种通信装置，包括至少一个处理器，所述至少一个处理器与存储器耦合；该存储器用于存储程序或指令；该至少一个处理器用于执行该程序或指令，以使该装置实现前述第三方面或第三方面任意一种可能的实现方式所述的方法。

本申请实施例第十六方面提供了一种通信装置，包括至少一个处理器，所述至少一个处理器与存储器耦合；该存储器用于存储程序或指令；该至少一个处理器用于执行该程序或指令，以使该装置实现前述第四方面或第四方面任意一种可能的实现方式所述的方法。

本申请实施例第十七方面提供了一种通信装置，包括至少一个处理器，所述至少一个处理器与存储器耦合；该存储器用于存储程序或指令；该至少一个处理器用于执行该程序或指令，以使该装置实现前述第五方面或第五方面任意一种可能的实现方式所述的方法。

本申请实施例第十八方面提供了一种通信装置，包括至少一个处理器，所述至少一个处理器与存储器耦合；该存储器用于存储程序或指令；该至少一个处理器用于执行该程序或指令，以使该装置实现前述第六方面或第六方面任意一种可能的实现方式所述的方法。

本申请实施例第十九方面提供了一种通信装置，包括至少一个逻辑电路和输入输出接口；该逻辑电路用于执行如前述第一方面及其任意一种可能的实现方式所述的方法。

本申请实施例第二十方面提供了一种通信装置，包括至少一个逻辑电路和输入输出接口；该逻辑电路用于执行如前述第二方面及其任意一种可能的实现方式所述的方法。

本申请实施例第二十一方面提供了一种通信装置，包括至少一个逻辑电路和输入输出接口；该逻辑电路用于执行如前述第三方面及其任意一种可能的实现方式所述的方法。

本申请实施例第二十二方面提供了一种通信装置，包括至少一个逻辑电路和输入输出接口；该逻辑电路用于执行如前述第四方面及其任意一种可能的实现方式所述的方法。

本申请实施例第二十三方面提供了一种通信装置，包括至少一个逻辑电路和输入输出接口；该逻辑电路用于执行如前述第五方面及其任意一种可能的实现方式所述的方法。

本申请实施例第二十四方面提供了一种通信装置，包括至少一个逻辑电路和输入输出接口；该逻辑电路用于执行如前述第六方面及其任意一种可能的实现方式所述的方法。

本申请实施例第二十五方面提供一种计算机可读存储介质，该存储介质用于存储一个或多个计算机执行指令，当计算机执行指令被处理器执行时，该处理器执行如上述第一方面至第六方面中任一方面的任意一种可能的实现方式所述的方法。

本申请实施例第二十六方面提供一种计算机程序产品（或称计算机程序），当计算机程序产品被该处

理器执行时，该处理器执行上述第一方面至第六方面中任一方面的任意一种可能的实现方式所述的方法。

本申请实施例第二十七方面提供了一种芯片系统，该芯片系统包括至少一个处理器，用于支持通信装置实现上述第一方面至第六方面中任一方面的任意一种可能的实现方式中所涉及的功能。

在一种可能的设计中，该芯片系统还可以包括存储器，存储器，用于保存该第一通信装置必要的程序指令和数据。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。可选的，所述芯片系统还包括接口电路，所述接口电路为所述至少一个处理器提供程序指令和/或数据。

本申请实施例第二十八方面提供了一种通信系统，该通信系统包括上述第七方面的通信装置和第八方面的通信装置，和/或，该通信系统包括上述第九方面的通信装置和第十方面的通信装置，和/或，该通信系统包括上述第十一方面的通信装置和第十二方面的通信装置，和/或，该通信系统包括上述第十三方面的通信装置和第十四方面的通信装置，和/或，该通信系统包括上述第十五方面的通信装置和第十六方面的通信装置，和/或，该通信系统包括上述第十七方面的通信装置和第十八方面的通信装置，和/或，该通信系统包括上述第十九方面的通信装置和第二十方面的通信装置，和/或，该通信系统包括上述第二十一方面的通信装置和第二十二方面的通信装置，和/或，该通信系统包括上述第二十三方面的通信装置和第二十四方面的通信装置。

其中，第七方面至第二十八方面中任一种设计方式所带来的技术效果可参见上述第一方面至第六方面中不同设计方式所带来的技术效果，在此不再赘述。

附图说明

图1a为本申请提供的通信系统的一个示意图；
图1b为本申请提供的通信系统的另一个示意图；
图1c为本申请提供的通信系统的另一个示意图；
图1d为本申请涉及的AI处理过程的一个示意图；
图1e为本申请涉及的AI处理过程的另一个示意图；
图2a为本申请涉及的AI处理过程的另一个示意图；
图2b为本申请涉及的AI处理过程的另一个示意图；
图2c为本申请涉及的AI处理过程的另一个示意图；
图2d为本申请涉及的AI处理过程的另一个示意图；
图3为本申请提供的通信方法的一个交互示意图；
图4a为本申请提供的AI处理过程的另一个示意图；
图4b为本申请提供的AI处理过程的另一个示意图；
图4c为本申请提供的AI处理过程的另一个示意图；
图5为本申请提供的通信方法的另一个交互示意图；
图6为本申请提供的通信方法的另一个交互示意图；
图7为本申请提供的AI处理过程的另一个示意图；
图8为本申请提供的通信装置的一个示意图；
图9为本申请提供的通信装置的另一个示意图；
图10为本申请提供的通信装置的另一个示意图；
图11为本申请提供的通信装置的另一个示意图；
图12为本申请提供的通信装置的另一个示意图。

具体实施方式

首先，对本申请实施例中的部分用语进行解释说明，以便于本领域技术人员理解。
(1) 终端设备：可以是能够接收网络设备调度和指示信息的无线终端设备，无线终端设备可以是指向用户提供语音和/或数据连通性的设备，或具有无线连接功能的手持式设备，或连接到无线调制解调器的其他处理设备。

终端设备可以经无线接入网（radio access network, RAN）与一个或多个核心网或者互联网进行通信，终端设备可以是移动终端设备，如移动电话（或称为“蜂窝”电话，手机（mobile phone））、计算机和数据卡，例如，可以是便携式、袖珍式、手持式、计算机内置的或者车载的移动装置，它们与无线接入网交换语音和/或数据。例如，个人通信业务（personal communication service, PCS）电话、无绳电话、
5 会话发起协议（SIP）话机、无线本地环路（wireless local loop, WLL）站、个人数字助理（personal digital assistant, PDA）、平板电脑（Pad）、带无线收发功能的电脑等设备。无线终端设备也可以称为系统、订户单元（subscriber unit）、订户站（subscriber station）、移动站（mobile station）、移动台（mobile station, MS）、远程站（remote station）、接入点（access point, AP）、远程终端设备（remote terminal）、接入终端设备（access terminal）、用户终端设备（user terminal）、用户代理（user agent）、
10 用户站（subscriber station, SS）、用户端设备（customer premises equipment, CPE）、终端（terminal）、用户设备（user equipment, UE）、移动终端（mobile terminal, MT）等。

作为示例而非限定，在本申请实施例中，该终端设备还可以是可穿戴设备。可穿戴设备也可以称为穿戴式智能设备或智能穿戴式设备等，是应用穿戴式技术对日常穿戴进行智能化设计、开发出可以穿戴的设备的总称，如眼镜、手套、手表、服饰及鞋等。可穿戴设备即直接穿在身上，或是整合到用户的衣服或配件的一种便携式设备。可穿戴设备不仅仅是一种硬件设备，更是通过软件支持以及数据交互、云端交互来实现强大的功能。广义穿戴式智能设备包括功能全、尺寸大、可不依赖智能手机实现完整或者部分的功能，例如：智能手表或智能眼镜等，以及只专注于某一类应用功能，需要和其它设备如智能手机配合使用，如各类进行体征监测的智能手环、智能头盔、智能首饰等。

终端还可以是无人机、机器人、设备到设备通信（device-to-device, D2D）中的终端、车到一切（vehicle to everything, V2X）中的终端、虚拟现实（virtual reality, VR）终端设备、增强现实（augmented reality, AR）终端设备、工业控制（industrial control）中的无线终端、无人驾驶（self driving）中的无线终端、远程医疗（remote medical）中的无线终端、智能电网（smart grid）中的无线终端、运输安全（transportation safety）中的无线终端、智慧城市（smart city）中的无线终端、智慧家庭（smart home）中的无线终端等。

此外，终端设备也可以是第五代（5th generation, 5G）通信系统之后演进的通信系统（例如第六代（6th generation, 6G）通信系统等）中的终端设备或者未来演进的公共陆地移动网络（public land mobile network, PLMN）中的终端设备等。示例性的，6G网络可以进一步扩展5G通信终端的形态和功能，6G终端包括但不限于车、蜂窝网络终端(融合卫星终端功能)、无人机、物联网（internet of things, IoT）设备。

在本申请实施例中，上述终端设备还可以获得网络设备提供的AI服务。可选地，终端设备还可以具有AI处理能力。

（2）网络设备：可以是无线网络中的设备，例如网络设备可以为将终端设备接入到无线网络的RAN节点（或设备），又可以称为基站。目前，一些RAN设备的举例为：基站（base station）、演进型基站（evolved NodeB, eNodeB）、5G通信系统中的基站gNB（gNodeB）、传输接收点（transmission reception point, TRP）、演进型节点B（evolved Node B, eNB）、无线网络控制器（radio network controller, RNC）、节点B（Node B, NB）、家庭基站（例如，home evolved Node B, 或home Node B, HNB）、基带单元（base band unit, BBU），或无线保真（wireless fidelity, Wi-Fi）接入点AP等。另外，在一种网络结构中，网络设备可以包括集中单元（centralized unit, CU）节点、或分布单元（distributed unit, DU）节点、或包括CU节点和DU节点的RAN设备。

可选的，RAN节点还可以是宏基站、微基站或室内站、中继节点或施主节点、或者是云无线接入网络（cloud radio access network, CRAN）场景下的无线控制器。RAN节点还可以是服务器，可穿戴设备，车辆或车载设备等。例如，车辆外联（vehicle to everything, V2X）技术中的接入网设备可以为路侧单元（road side unit, RSU）。

在另一种可能的场景中，由多个RAN节点协作协助终端实现无线接入，不同RAN节点分别实现基站的部分功能。例如，RAN节点可以是集中式单元（central unit, CU），分布式单元（distributed unit, DU），CU-控制面（control plane, CP），CU-用户面（user plane, UP），或者无线单元（radio unit, RU）等。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。

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, 在开放式接入网(open RAN, O-RAN或ORAN)系统中, CU也可以称为O-CU(开
5 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。

接入网设备和终端设备之间的通信遵循一定的协议层结构。该协议层可以包括控制面协议层和用户面
10 协议层。控制面协议层可以包括以下至少一项: 无线资源控制(radio resource control, RRC)层、分组数据汇聚层协议(packet data convergence protocol, PDCP)层、无线链路控制(radio link control, RLC)层、媒体接入控制(media access control, MAC)层、或物理(physical, PHY)层等。用户面协议层可以包括以下至少一项: 业务数据适配协议(service data adaptation protocol, SDAP)层、PDCP层、RLC层、MAC层、或物理层等。

15 对于ORAN系统中的网元及其可实现的协议层功能对应关系, 可参照下表1。

表1

ORAN网元	3GPP的协议层功能
O-CU-CP	RRC+PCDP-控制面(PDCP-C)
O-CU-UP	SDAP+PCDP-用户面(PDCP-U)
O-DU	RLC+MAC+PHY-high
O-RU	PHY-low

网络设备可以是其它为终端设备提供无线通信功能的装置。本申请的实施例对网络设备所采用的具体
20 技术和具体设备形态不做限定。为方便描述, 本申请实施例并不限定。

网络设备还可以包括核心网设备, 核心网设备例如包括第四代(4th generation, 4G)网络中的移动性
管理实体(mobility management entity, MME), 归属用户服务器(home subscriber server, HSS), 服务
网关(serving gateway, S-GW), 策略和计费规则功能(policy and charging rules function, PCRF), 公共
数据网网关(public data network gateway, PDN gateway, P-GW); 5G网络中的访问和移动管理功能
25 (access and mobility management function, AMF)、用户面功能(user plane function, UPF)或会话管理
功能(session management function, SMF)等网元。此外, 该核心网设备还可以包括5G网络以及5G网络的
下一代网络中的其他核心网设备。

本申请实施例中, 上述网络设备还可以具有AI能力的网络节点, 可以为终端或其他网络设备提供AI服
务, 例如, 可以为网络侧(接入网或核心网)的AI节点、算力节点、具有AI能力的RAN节点、具有AI能力
30 的核心网网元等。

本申请实施例中, 用于实现网络设备的功能的装置可以是网络设备, 也可以是能够支持网络设备实现
该功能的装置, 例如芯片系统, 该装置可以被安装在网络设备中。在本申请实施例提供的技术方案中, 以
用于实现网络设备的功能的装置是网络设备为例, 描述本申请实施例提供的技术方案。

(3) 配置与预配置: 在本申请中, 会同时用到配置与预配置。其中, 配置是指网络设备/服务器通过
35 消息或信令将一些参数的配置信息或参数的取值发送给终端, 以便终端根据这些取值或信息来确定通信的
参数或传输时的资源。预配置与配置类似, 可以是网络设备/服务器预先与终端设备协商好的参数信息或参
数值, 也可以是标准协议规定的基站/网络设备或终端设备采用的参数信息或参数值, 还可以是预先存储在
基站/服务器或终端设备的参数信息或参数值。本申请对此不做限定。

进一步地, 这些取值和参数, 是可以变化或更新的。

40 (4) 本申请实施例中的术语“系统”和“网络”可被互换使用。“多个”是指两个或两个以上。
“和/或”, 描述关联对象的关联关系, 表示可以存在三种关系, 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的情况，其中A，B可以是单数或者复数。字符“/”一般表示前后关联对象是一种“或”的关系。“以下至少一项（个）”或其类似表达，是指的这些项中的任意组合，包括单项（个）或复数项（个）的任意组合。例如“A，B和C中的至少一项”包括A，B，C，AB，AC，BC或ABC。以及，除非有特别说明，本申请实施例提及“第一”、“第二”等序数词是用于对多个对象进行区分，不用于限定多个对象的顺序、时序、优先级或者重要程度。

（5）本申请实施例中的“发送”和“接收”，表示信号传递的走向。例如，“向XX发送信息”可以理解为该信息的目的端是XX，可以包括通过空口直接发送，也包括其他单元或模块通过空口间接发送。

“接收来自YY的信息”可以理解为该信息的源端是YY，可以包括通过空口直接从YY接收，也可以包括通过空口从其他单元或模块间接地从YY接收。“发送”也可以理解为芯片接口的“输出”，“接收”也可以理解为芯片接口的“输入”。

换言之，发送和接收可以是在设备之间进行的，例如，网络设备和终端设备之间进行的，也可以是在设备内进行的，例如，通过总线、走线或接口在设备内的部件之间、模组之间、芯片之间、软件模块或者硬件模块之间发送或接收。

可以理解的是，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，比如编码、调制等，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相似的理解，不再赘述。

（6）在本申请实施例中，“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。将某一信息（如下文所述的指示信息）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，例如可以借助预先约定（例如协议预定义）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。本申请对于指示的具体方式不作限定。可以理解的是，对于该指示信息的发送方来说，该指示信息可用于指示待指示信息，对于指示信息的接收方来说，该指示信息可用于确定待指示信息。

本申请中，除特别说明外，各个实施例之间相同或相似的部分可以互相参考。在本申请中各个实施例、以及各实施例中的各个方法/设计/实现方式中，如果没有特殊说明以及逻辑冲突，不同的实施例之间、以及各实施例中的各个方法/设计/实现方式之间的术语和/或描述具有一致性、且可以相互引用，不同的实施例、以及各实施例中的各个方法/设计/实现方式中的技术特征根据其内在的逻辑关系可以组合形成新的实施例、方法、或实现方式。以下所述的本申请实施方式并不构成对本申请保护范围的限定。

本申请可以应用于长期演进（long term evolution，LTE）系统、新无线（new radio，NR）系统，或者是5G之后演进的通信系统（例如6G等）。其中，该通信系统中包括至少一个网络设备和/或至少一个终端设备。

请参阅图1a，为本申请中通信系统的一种示意图。图1a中，示例性的示出了一个网络设备和6个终端设备，6个终端设备分别为终端设备1、终端设备2、终端设备3、终端设备4、终端设备5以及终端设备6等。

在图1a所示的示例中，是以终端设备1为智能茶杯，终端设备2为智能空调，终端设备3为智能加油机，终端设备4为交通工具，终端设备5为手机，终端设备6为打印机进行举例说明的。

如图1a所示，AI配置信息发送实体可以为网络设备。AI配置信息接收实体可以为终端设备1-终端设备6，此时，网络设备和终端设备1-终端设备6组成一个通信系统，在该通信系统中，终端设备1-终端设备6可以发送数据给网络设备，网络设备需要接收终端设备1-终端设备6发送的数据。同时，网络设备可以向终端设备1-终端设备6发送配置信息。

示例性的，在图1a中，终端设备4-终端设备6也可以组成一个通信系统。其中，终端设备5作为网络设备，即AI配置信息发送实体；终端设备4和终端设备6作为终端设备，即AI配置信息接收实体。例如车联网系统中，终端设备5分别向终端设备4和终端设备6发送AI配置信息，并且接收终端设备4和终端设备6发送的数据；相应的，终端设备4和终端设备6接收终端设备5发送的AI配置信息，并向终端设备5发送数据。

以图1a所示通信系统为例，不同的设备之间（包括网络设备与网络设备之间，网络设备与终端设备之间，和/或，终端设备和终端设备之间）除了执行通信相关业务之外，还有可能执行AI相关业务。例如，如

图1b所示，以网络设备为基站为例，基站可以与一个或多个终端设备之间可以执行通信相关业务和AI相关业务，不同终端设备之间也可以执行通信相关业务和AI相关业务。又如，如图1c所示，以终端设备包括电视和手机为例，电视和手机之间也可以执行通信相关业务和AI相关业务。

本申请提供的技术方案可以应用于无线通信系统（例如图 1a、图 1b 或图 1c 所示系统），例如本申请提供的通信系统中可以引入 AI 网元来实现部分或全部 AI 相关的操作。AI 网元也可以称为 AI 节点、AI 设备、AI 实体、AI 模块、AI 模型、或 AI 单元等。所述 AI 网元可以是内置在通信系统的网元中。例如，AI 网元可以是内置在：接入网设备、核心网设备、云服务器、或网管（operation, administration and maintenance, OAM）中的 AI 模块，用以实现 AI 相关的功能。所述 OAM 可以是作为核心网设备网管和/或作为接入网设备的网管。或者，所述 AI 网元也可以是通信系统中独立设置的网元。可选的，终端或终端内置的芯片中也可以包括 AI 实体，用于实现 AI 相关的功能。

下面将本申请中可能涉及到的人工智能（artificial intelligence, AI）进行简要介绍。

人工智能（artificial intelligence, AI），可以让机器具有人类的智能，例如可以让机器应用计算机的软硬件来模拟人类某些智能行为。为了实现人工智能，可以采用机器学习方法。机器学习方法中，机器利用训练数据学习（或训练）得到模型。该模型表征了从输入到输出之间的映射。学习得到的模型可以用于进行推理（或预测），即可以利用该模型预测出给定输入所对应的输出。其中，该输出还可以称为推理结果（或预测结果）。

机器学习可以包括监督学习、无监督学习、和强化学习。其中，无监督学习还可以称为非监督学习。

监督学习依据已采集到的样本值和样本标签，利用机器学习算法学习样本值到样本标签的映射关系，并用 AI 模型来表达学到的映射关系。训练机器学习模型的过程就是学习这种映射关系的过程。在训练过程中，将样本值输入模型得到模型的预测值，通过计算模型的预测值与样本标签（理想值）之间的误差来优化模型参数。映射关系学习完成后，就可以利用学到的映射来预测新的样本标签。监督学习学到的映射关系可以包括线性映射或非线性映射。根据标签的类型可将学习的任务分为分类任务和回归任务。

无监督学习依据采集到的样本值，利用算法自行发掘样本的内在模式。无监督学习中有一类算法将样本自身作为监督信号，即模型学习从样本到样本的映射关系，称为自监督学习。训练时，通过计算模型的预测值与样本本身之间的误差来优化模型参数。自监督学习可用于信号压缩及解压恢复的应用，常见的算法包括自编码器和对抗生成网络等。

强化学习不同于监督学习，是一类通过与环境进行交互来学习解决问题的策略的算法。与监督、无监督学习不同，强化学习问题并没有明确的“正确的”动作标签数据，算法需要与环境进行交互，获取环境反馈的奖励信号，进而调整决策动作以获得更大的奖励信号数值。如下行功率控制中，强化学习模型根据无线网络反馈的系统总吞吐率，调整各个用户的下行发送功率，进而期望获得更高的系统吞吐率。强化学习的目标也是学习环境状态与较优（例如最优）决策动作之间的映射关系。但因为无法事先获得“正确动作”的标签，所以不能通过计算动作与“正确动作”之间的误差来优化网络。强化学习的训练是通过与环境的迭代交互而实现的。

神经网络（neural network, NN）是机器学习技术中的一种具体的模型。根据通用近似定理，神经网络在理论上可以逼近任意连续函数，从而使得神经网络具备学习任意映射的能力。传统的通信系统需要借助丰富的专家知识来设计通信模块，而基于神经网络的深度学习通信系统可以从大量的数据集中自动发现隐含的模式结构，建立数据之间的映射关系，获得优于传统建模方法的性能。

神经网络的思想来源于大脑组织的神经元结构。例如，每个神经元都对其输入值进行加权求和运算，通过一个激活函数输出运算结果。如图 1d 所示，为神经元结构的一种示意图。假设神经元的输入为 $x = [x_0, x_1, \dots, x_n]$ ，与各个输入对应的权值分别为 $w = [w, w_1, \dots, w_n]$ ，其中， n 为正整数， w_i 和 x_i 可以是小数、整数（例如 0、正整数或负整数等）、或复数等各种可能的类型。 w_i 作为 x_i 的权值，用于对 x_i 进行加权。根据权值对输入值进行加权求和的偏置例如为 b 。激活函数的形式可以有多种，假设一个神经元的激活函数为： $y = f(z) = \max(0, z)$ ，则该神经元的输出为： $y = f(\sum_{i=0}^n w_i * x_i + b) = \max(0, \sum_{i=0}^n w_i * x_i + b)$ 。再例如，一个神经元的激活函数为： $y = f(z) = z$ ，则该神经元的输出为： $y = f(\sum_{i=0}^n w_i * x_i + b) =$

$\sum_{i=0}^{i=n} w_i * x_i + b$ 。其中，b 可以是小数、整数（例如 0、正整数或负整数）、或复数等各种可能的类型。神经网络中不同神经元的激活函数可以相同或不同。

此外，神经网络一般包括多个层，每层可包括一个或多个神经元。通过增加神经网络的深度和/或宽度，能够提高该神经网络的表达能力，为复杂系统提供更强大的信息提取和抽象建模能力。其中，神经网络的深度可以是指神经网络包括的层数，每层包括的神经元个数可以称为该层的宽度。在一种实现方式中，神经网络包括输入层和输出层。神经网络的输入层将接收到的输入信息经过神经元处理，将处理结果传递给输出层，由输出层得到神经网络的输出结果。在另一种实现方式中，神经网络包括输入层、隐藏层和输出层。神经网络的输入层将接收到的输入信息经过神经元处理，将处理结果传递给中间的隐藏层，隐藏层对接收的处理结果进行计算，得到计算结果，隐藏层将计算结果传递给输出层或者下一个相邻的隐藏层，最终由输出层得到神经网络的输出结果。其中，一个神经网络可以包括一个隐藏层，或者包括多个依次连接的隐藏层，不予限制。

神经网络例如为深度神经网络（deep neural network, DNN）。根据网络的构建方式，DNN 可以包括前馈神经网络（feedforward neural network, FNN）、卷积神经网络（convolutional neural networks, CNN）和递归神经网络（recurrent neural network, RNN）。

FNN 网络的特点为相邻层的神经元之间两两完全相连。该特点使得 FNN 通常需要大量的存储空间、导致较高的计算复杂度。图 1e 为一种 FNN 网络示意图。

CNN 是一种专门来处理具有类似网格结构的数据的神经网络。例如，时间序列数据（时间轴离散采样）和图像数据（二维离散采样）都可以认为是类似网格结构的数据。CNN 并不一次性利用全部的输入信息做运算，而是采用一个固定大小的窗截取部分信息做卷积运算，这就大大降低了模型参数的计算量。另外根据窗截取的信息类型的不同（如同一副图中的人和物为不同类型信息），每个窗可以采用不同的卷积核运算，这使得 CNN 能更好的提取输入数据的特征。

RNN 是一类利用反馈时间序列信息的 DNN 网络。它的输入包括当前时刻的新的输入值和自身在前一时刻的输出值。RNN 适合获取在时间上具有相关性的序列特征，特别适用于语音识别、信道编译码等应用。

在上述机器学习的模型训练过程中，可以定义损失函数。损失函数描述了模型的输出值和理想目标值之间的差距或差异。损失函数可以通过多种形式体现，对于损失函数的具体形式不予限制。模型训练过程可以看作以下过程：通过调整模型的部分或全部参数，使得损失函数的值小于门限值或者满足目标需求。

模型还可以被称为 AI 模型、规则或者其他名称等。AI 模型可以认为是实现 AI 功能的具体方法。AI 模型表征了模型的输入和输出之间的映射关系或者函数。AI 功能可以包括以下一项或多项：数据收集、模型训练（或模型学习）、模型信息发布、模型推断（或称为模型推理、推理、或预测等）、模型监控或模型校验、或推理结果发布等。AI 功能还可以称为 AI（相关的）操作、或 AI 相关的功能。

下面将结合附图，对神经网络的实现过程进行示例性描述。

1. 全连接神经网络。

又叫多层感知机（multilayer perceptron, MLP）。如图 2a 所示，一个 MLP 包含一个输入层（左侧），一个输出层（右侧），及多个隐藏层（中间）。其中，MLP 的每层包含若干个节点，称为神经元。其中，相邻两层的神经元间两两相连。

可选的，考虑相邻两层的神经元，下一层的神经元的输出 h 为所有与之相连的上一层神经元 x 的加权和并经过激活函数，可以表示为：

$$h = f(wx + b)。$$

其中，w 为权重矩阵，b 为偏置向量，f 为激活函数。

进一步可选的，神经网络的输出可以递归表达为：

$$y = f_n(w_n f_{n-1}(\dots) + b_n)。$$

其中，n 是神经网络层的索引， $1 \leq n \leq N$ ，其中 N 为神经网络的总层数。

换言之，可以将神经网络理解为一个从输入数据集到输出数据集的映射关系。而通常神经网络都是随机初始化的，用已有数据从随机的 w 和 b 得到这个映射关系的过程被称为神经网络的训练。

可选的，训练的具体方式为采用损失函数（loss function）对神经网络的输出结果进行评价。如图2b所示，可以将误差反向传播，通过梯度下降的方法即能迭代优化神经网络参数（包括 $\mathbf{w}$ 和 $\mathbf{b}$ ），直到损失函数达到最小值，即图2b中的“较优点（例如最优值）”。可以理解的是，图2b中的“较优点（例如最优值）”对应的神经网络参数可以作为训练好的AI模型信息中的神经网络参数。

进一步可选的，梯度下降的过程可以表示为：

$$\theta \leftarrow \theta - \eta \frac{\partial L}{\partial \theta}$$

其中， θ 为待优化参数（包括 $\mathbf{w}$ 和 $\mathbf{b}$ ）， L 为损失函数， η 为学习率，控制梯度下降的步长， ∂ 表示求导运算， $\frac{\partial L}{\partial \theta}$ 表示对 L 求 θ 的导数。

进一步可选的，反向传播的过程利用到求偏导的链式法则。如图2c所示，前一层参数的梯度可以由后一层参数的梯度递推计算得到，可以表达为：

$$\frac{\partial L}{\partial w_{ij}} = \frac{\partial L}{\partial s_i} \frac{\partial s_i}{\partial w_{ij}}$$

其中， w_{ij} 为节点 j 连接节点 i 的权重， s_i 为节点 i 上的输入加权和。

2. 联邦学习（Federated Learning, FL）。

联邦学习这一概念的提出有效地解决了当前人工智能发展所面临的困境，其在充分保障用户数据隐私和安全的前提下，通过促使各个边缘设备和中心端服务器协同合作来高效地完成模型的学习任务。如图2d所示，FL架构是当前FL领域最为广泛的训练架构，FedAvg算法是FL的基础算法，其算法流程大致如下：

（1）中心端初始化待训练模型 $\mathbf{w}_g^0$ 并将其广播发送给所有客户端设备。

（2）在第 $t \in [1, T]$ 轮中，客户端 $k \in [1, K]$ 基于局部数据集 $\mathcal{D}_k$ 对接收到的全局模型 $\mathbf{w}_g^{t-1}$ 进行 E 个epoch的训练以得到本地训练结果 $\mathbf{w}_k^t$ ，将其上报给中心节点。

（3）中心节点汇总收集来自全部（或部分）客户端的本地训练结果，假设第 t 轮上传局部模型的客户端集合为 $\mathcal{S}^t$ ，中心端将以对应客户端的样本数为权重进行加权求均得到新的全局模型，具体更新法则为

$$\mathbf{w}_g^t = \sum_{k \in \mathcal{S}^t} \frac{D_k \mathbf{w}_k^t}{\sum_{k \in \mathcal{S}^t} D_k}$$

其后中心端再将最新版本的全局模型 $\mathbf{w}_g^t$ 广播发送给所有客户端设备进行新一轮的训练。

（4）重复步骤（2）和（3）直至模型最终收敛或训练轮数达到上限。

除了上报本地模型 $\mathbf{w}_k^t$ ，还可以将训练的本地梯度 $\mathbf{g}_k^t$ 进行上报，中心节点将本地梯度求平均，并根据这个平均梯度的方向更新全局模型。

可以看到，在FL框架中，数据集存在于分布式节点处，即分布式节点收集本地的数据集，并进行本地训练，将训练得到的本地结果（模型或梯度）上报给中心节点。中心节点本身没有数据集，只负责将分布式节点的训练结果进行融合处理，得到全局模型，并下发给分布式节点。

3. 知识蒸馏。

异构模型融合的典型方法是通过知识蒸馏的方式，将不同结构的模型转换至相同结构上，而后进行融合。所谓知识蒸馏，是将一个网络的知识转移到另一个网络，两个网络可以是同构或者异构。知识蒸馏，可以用来将网络从大网络转化成一个网络（例如，大网络的参数规模大于小网络的参数规模），并保留接近于大网络的性能；也可以将多个网络的学到的知识转移到一个网络中。异构融合的具体做法是中心节点先收集各个分布节点的异构模型，将其蒸馏至同构的模型、并融合，再将融合后的同构模型反馈至各个分布式节点，并由分布式节点恢复到自身的网络结构，开启下一轮融合。

本申请提供的技术方案可以应用于无线通信系统（例如图1a或图1b所示系统），在无线通信系统中，通信节点一般具备信号收发能力和计算能力。以具备计算能力的网络设备为例，网络设备的计算能力主要是为信号收发能力提供算力支持（例如：对承载信号的时域资源、频域资源等进行计算），以实现网络设备与其它通信节点的通信任务。

然而，在通信网络中，通信节点的计算能力除了为上述通信任务提供算力支持之外，还可能具备富余的计算能力。为此，如何利用这些计算能力，是一个亟待解决的技术问题。

作为一种实现示例，为了应对未来智能普惠的愿景，智能化有可能将在无线网络架构层面进一步演进，在未来的通信系统中，AI将有可能与无线网络进一步深度的融合，实现网络内生的智能，同时还包括终端的智能化。示例性的，AI与无线网络的融合可以应用于如下的新需求和新场景：例如，终端连接更加灵活和智能：包括但不限于终端类型多样化，超级物联网（Supper IOT）（例如Supper IOT可以包括物联，车联，工业，医疗...），海量连接，终端连接更加灵活，终端本身具备一定的AI能力。又如：网络内生智能：网络除了提供传统的通信连接服务，还将进一步提供计算和AI服务，来更好的支持普惠性的、实时性的和高安全的AI服务。这些新需求和新场景会带来无线网络架构和通信模式的变化。

一般地，AI学习的参与节点可以包括分布式的多个通信网络中的节点，例如终端设备、网络设备等。例如，在图2d所示存在中心节点的网络架构中，中心节点可以为网络设备或终端设备，任一分布式节点也可以为网络设备或终端设备。在AI学习过程中，不同节点可以基于随机化的神经网络模型参与AI学习，即各个节点获取随机化的模型参数之后，各个节点均可以在本地基于该随机化的模型参数对神经网络模型进行初始化处理得到初始化的神经网络模型，并基于该初始化的神经网络模型执行AI学习过程，直至AI神经网络模型收敛至预设条件（例如图2b中神经网络模型的损失函数指示该模型收敛至较优点（例如最优点），又如神经网络模型的迭代次数达到阈值等）。

然而，不同节点的本地信息有可能是不同的，例如不同节点的计算资源（例如CPU、GPU等模块/器件的资源）、不同节点的存储资源（例如ROM、RAM等模块/器件的资源）、以及不同节点的本地的计算资源的利用率中的一项或多项有可能是不同的。为此，上述各个节点基于随机化的模型参数进行AI学习的过程中，各个节点所获得的初始化的神经网络模型是随机化的，这将有可能会导致AI神经网络模型收敛的速度变慢，进而将会影响AI学习的处理效率。

为了解决上述问题，本申请提供了一种通信方法及相关设备，用于使得通信节点的算力能够应用于人工智能（artificial intelligence, AI）学习，并能够提升不同节点的实现灵活性。

请参阅图3，为本申请提供的通信方法的一个实现示意图，该方法包括如下步骤。

需要说明的是，图3中以第一节点和第二节点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，图3及对应的实现方式中S301的执行主体是第一节点，执行主体也可以是支持该第一节点实现该方法的芯片、芯片系统、或处理器，还可以是能实现全部或部分第一节点功能的逻辑模块或软件。图3及对应的实现方式中S301-S302中的第二节点也可以替换为支持该第二节点实现该方法的芯片、芯片系统、或处理器，还可以替换为能实现全部或部分第二节点功能的逻辑模块或软件。

S301.第二节点发送第一信息，相应的，第一节点接收该第一信息。其中，该第一信息用于指示N个配置信息，该N个配置信息用于配置N个模型参数的资源，N为正整数。

需要说明的是，不同的第一节点可以为分布式节点，第二节点可以与一个或多个第一节点连接。其中，第二节点可以作为该一个或多个第一节点的中心节点（或控制节点，管理节点等）；或者，第二节点可以作为服务端（server），且一个或多个第一节点作为客户端（client）。

应理解，第一信息用于指示N个配置信息，可以理解为：第一信息包括该N个配置信息的索引，使得该第一信息的接收方能够基于该索引获得该N个配置信息；或者，第一信息包括该N个配置信息，使得该第一信息的接收方能够从该第一信息中获得该N个配置信息。此外，N个配置信息用于配置N个模型参数的资源，可以理解为N个配置信息分别用于配置N个模型参数的资源，或者，N个配置信息与N个模型参数的资源一一对应。此外，第二节点可以通过一个或多个信息/消息发送给N个配置信息，即第一信息指示的N

个配置信息可以一次发送的过程完成，也可以通过多次发送的过程完成。相应的，第一节点也可以通过一次接收的过程接收该N个配置信息，也可以通过多次接收的过程接收该N个配置信息。

类似地，后文提及的第二信息用于指示M个配置信息，可以理解为：第二信息包括该M个配置信息的索引，使得该第二信息的接收方能够基于该索引获得该M个配置信息；或者，第二信息包括该M个配置信息，使得该第二信息的接收方能够从该M个配置信息。此外，M个配置信息用于配置M个模型参数的资源，可以理解为M个配置信息分别用于配置M个模型参数的资源，或者，M个配置信息与M个模型参数的资源一一对应。此外，第二节点可以通过一个或多个信息/消息发送给N个配置信息，即第二信息指示的M个配置信息可以一次发送的过程完成，也可以通过多次发送的过程完成。相应的，第一节点也可以通过一次接收的过程接收该M个配置信息，也可以通过多次接收的过程接收该M个配置信息。

类似地，后文提及的第三信息用于指示K个配置信息，可以理解为：第三信息包括该K个配置信息的索引，使得该第三信息的接收方能够基于该索引获得该K个配置信息；或者，第三信息包括该K个配置信息，使得该第三信息的接收方能够从该K个配置信息。此外，K个配置信息用于配置K个模型参数的资源，可以理解为K个配置信息分别用于配置K个模型参数的资源，或者，K个配置信息与K个模型参数的资源一一对应。此外，第二节点可以通过一个或多个信息/消息发送给N个配置信息，即第三信息指示的K个配置信息可以一次发送的过程完成，也可以通过多次发送的过程完成。相应的，第一节点也可以通过一次接收的过程接收该K个配置信息，也可以通过多次接收的过程接收该K个配置信息。

在一种可能的实现方式中，在步骤S301中，对于第二节点发送的第一信息指示的N个配置信息所配置的任一模型参数而言，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。从而，N个配置信息所配置的资源传输的N个模型参数包括上述至少一项，以提升方案实现的灵活性。

S302.第一节点基于N个配置信息中的第一配置信息接收第一模型参数，该第一模型参数用于确定与本地信息对应的第一神经网络模型。

可选地，AI学习可以替换为机器学习，AI处理，AI神经网络处理等。

应理解，N个配置信息所配置的资源包括时域资源、频域资源、空域资源中的一项或多项。在步骤S302中，第一节点基于N个配置信息中的第一配置信息接收第一模型参数可以理解为，该第一节点基于N个配置信息中的第一配置信息所配置的资源接收该第一模型参数。

应理解，在步骤S301中，第一信息指示的N个配置信息配置的N个模型参数中，模型参数用于神经网络模型的初始化，或者说，模型参数包括神经网络模型的初始化参数。其中，初始化参数可以是用于确定神经网络模型使用的参数，或，初始化参数可以是用于确定神经网络模型训练启动的参数。换言之，对于第一节点而言，该第一节点在步骤S301中基于第一配置信息接收第一模型参数之后，在步骤S302中，该第一节点可以基于该第一模型参数进行神经网络模型的初始化处理，以得到与本地信息对应的第一神经网络模型；或者，在步骤S302中，该第一节点可以将该第一模型参数作为神经网络模型的初始化参数，完成本地参数初始化之后，得到第一神经网络模型。

在一种可能的实现方式中，第一节点在步骤S302中接收的第一模型参数用于确定与本地信息对应的第一神经网络模型，该本地信息包括本地资源开销水平，本地AI能力信息、本地的参数规模中的至少一项。从而，本地信息可以通过多种方式实现，以提供方案实现的灵活性。

应理解，本地资源开销水平可以理解为本地的计算资源的利用率，本地的存储资源的利用率，本地功耗的使用情况等。例如，该本地的计算资源可以包括中央处理器（central processor unit, CPU）、微控制器（micro controller unit, MCU）或图形处理器（graphics processing unit, GPU）等模块/器件的资源；该本地的存储资源可以包括只读存储器（read only memory, ROM）、随机存取存储器（random access memory, RAM）等模块/器件的资源；该本地功耗的使用情况可以包括电池模块的健康度，电池模块当前已使用的电量（或电量的百分比），电池模块当前尚未使用的电量（或电量的百分比）等。

应理解，本地AI能力信息可以包括本地与AI相关的能力信息，例如与AI相关的计算能力、存储能力等。类似的，与AI相关的计算能力可以包括CPU、MCU、GPU等模块/器件的计算能力，与AI相关的存储能力可以包括ROM、RAM等模块/器件的存储能力。

应理解，本地的参数规模，可以理解为第一节点本地的神经网络模型的参数规模，包括但不限于神经网络模型的神经网络层数、神经网络模型的神经元数量，神经网络模型的输入输出或神经网络模型的计算复杂度不同等。

需要说明的是，第一节点的本地信息可以通过多种方式实现，相应的，对于第一节点而言，该第一节点接收的第一模型参数之后，该第一节点可以通过多种方式确定与本地信息对应的第一神经网络模型。

一种实现示例中，以该本地信息包括本地资源开销水平为例，该本地资源开销水平可以通过CPU利用率体现；其中，第一节点可以通过配置或预配置的方式确定不同的CPU利用率与不同的神经网络模型之间的映射关系。以不同的神经网络模型包括第一类神经网络模型和第二类神经网络模型为例，第一类神经网络模型可以比第二类神经网络模型更复杂（例如模型参数的规模更大，模型的神经网络层数更多等），该映射关系可以指示CPU利用率为0至X%（X取值为大于0且小于100）的节点对应第一类神经网络模型，并指示CPU利用率为X%至100%的节点对应第二类神经网络模型。从而，在上述技术方案中，第一节点可以基于本地的CPU利用率与X%进行比较，在本地的CPU利用率低于或等于X%的情况下，第一节点基于第一模型参数得到的第一神经网络模型可以为第一类神经网络模型；在本地的CPU利用率高于或等于X%的情况下，第一节点基于第一模型参数得到的第一神经网络模型可以为第二类神经网络模型。

可选地，在上述实现示例中，CPU利用率的测量可以是在节点执行相同或相似计算参数的这个前提下确定的，即在相对公平的情况下确定CPU利用率。比如，针对两个模型的训练，输入参数相同时，不同的节点为了完成一轮训练所需要的CPU利用率有差异。例如，在某个节点的CPU利用率较低的情况下，可以理解为该节点的CPU的算力较强，可以基于模型参数规模较大（或较复杂的模型参数）进行初始化处理；在某个节点的CPU利用率较高的情况下，可以理解为该节点的CPU的算力较弱，可以基于模型参数规模较小（或较简化的模型参数）进行初始化处理。相应的，基于该差异，可以为该不同的节点配置不同的模型参数，以使得各个节点能够获得差异化（或个性化）的初始化神经网络模型。

另一种实现示例中，以该本地信息包括本地AI能力信息为例，该本地AI能力信息可以通过RAM的存储能力体现；其中，第一节点可以通过配置或预配置的方式确定不同的RAM的存储能力与不同的神经网络模型之间的映射关系。以不同的神经网络模型包括第一类神经网络模型、第二类神经网络模型和第三类神经网络模型为例，第一类神经网络模型可以比第二类神经网络模型更复杂（例如模型参数的规模更大，模型的神经网络层数更多等），第二类神经网络模型可以比第三类神经网络模型更复杂；该映射关系可以指示RAM的存储能力为0至Y吉字节（Gigabyte，GB）（Y的取值为大于0）的节点对应第一类神经网络模型，并指示RAM的存储能力大于Y GB且小于或等于Z GB（Z的取值为大于Y）的节点对应第二类神经网络模型，并指示RAM的存储能力大于Z GB的节点对应第三类神经网络模型。从而，在上述技术方案中，第一节点可以基于本地的RAM的存储能力与Y和Z进行比较，在本地的RAM的存储能力低于或等于Y GB的情况下，第一节点基于第一模型参数得到的第一神经网络模型可以为第一类神经网络模型；在本地的RAM的存储能力大于Y GB且小于或等于Z GB的情况下，第一节点基于第一模型参数得到的第一神经网络模型可以为第二类神经网络模型，在本地的RAM的存储能力大于Z GB的情况下，第一节点基于第一模型参数得到的第一神经网络模型可以为第三类神经网络模型。

在一种可能的实现方式中，在步骤S301中，第二节点发送的第一信息指示的N个配置信息所配置的N个模型参数的参数规模不同，且N大于1。在步骤S302之前，该方法还包括：该第一节点基于该本地信息在该N个配置信息中确定该第一配置信息。具体地，在第一信息指示的N个配置信息大于1的情况下，该N个配置信息所配置的N个模型参数的参数规模不同，使得第一节点能够基于本地信息在N个配置信息中确定与该本地信息对应的第一配置信息，使得该第一节点能够基于该第一配置信息获得与该本地信息对应的模型参数。

应理解，N个模型参数的参数规模不同，可以理解为该N个模型参数对节点的本地信息的要求不同。例如，以该本地信息包括本地资源开销水平为例，N个模型参数中参数规模较大的模型参数要求本地资源开销水平低于阈值，以便于提供更多的未被使用的资源处理该参数规模较大的模型参数；又如，以该本地信息包括本地AI能力信息为例，N个模型参数中参数规模较大的模型参数要求本地AI能力信息高于阈值，以便于提供更多的AI资源处理该参数规模较大的模型参数。

可选地，N个模型参数的参数规模不同，可以替换为其它的描述，例如N个模型参数对应的神经网络层数不同，N个模型参数对应的神经元不同，N个模型参数对应的输入输出不同，N个模型参数的计算复杂度不同，或者，N个模型参数的计算能力要求不同，或者，N个模型参数的存储能力要求不同等。

作为一种实现示例，以第二节点为网络设备且第一节点为终端设备为例，该N个模型参数可以包括图4a中的三个模型参数。其中，这三个模型参数可以分别用于图4a中三类模型的初始化处理，这三类模型为图4a中的类型0模型（Cat 0 Model）、类型1模型（Cat 1 Model）和类型2模型（Cat 2 Model）。在图4a所示示例可以看出，不同的模型的神经网络层数不同，从而，终端设备可以基于本地部署的模型接收N个模型参数中的其中一个模型参数，例如，部署类型0模型的终端设备（即图中Cat0 终端设备）可以接收Cat 0 Model的模型参数，部署类型1模型的终端设备（即图中Cat1 终端设备）可以接收Cat 1 Model的模型参数，部署类型2模型的终端设备（即图中Cat2 终端设备）可以接收Cat 2 Model的模型参数，使得各个终端设备均能够接收与本地信息相匹配的模型参数，进而基于所接收的模型参数进行初始化处理，得到初始化的神经网络模型可以用于后续的AI学习（例如模型训练（Training））。

在一种可能的实现方式中，在步骤S301中，第二节点发送的第一信息指示的N个配置信息为一个配置信息（即第一配置信息，且N的取值为1）。其中，该第一模型参数用于配置至少两个第一节点共用的神经网络模型；在步骤S302之后，该方法还包括：该第一节点基于该本地信息将该至少两个第一节点共用的神经网络模型进行知识蒸馏处理，得到该第一神经网络模型。具体地，在第一信息指示的N个配置信息为一个配置信息（即第一配置信息）的情况下，该第一配置信息配置的第一模型参数用于配置至少两个第一节点共用的神经网络模型，使得该至少两个第一节点均能够基于该第一配置信息获得共用的神经网络模型，并基于各自的本地信息对该共用的神经网络模型进行知识蒸馏处理，以获得与该本地信息对应的第一神经网络模型。

应理解，至少两个第一节点共用的神经网络模型，可以理解为该至少两个第一节点能够识别/处理的神经网络模型。或者，至少两个第一节点共用的神经网络模型可以表述为：至少两个第一节点通用的神经网络模型，至少两个第一节点能够识别/处理的神经网络模型。

在一种可能的实现方式中，第一节点在步骤S301中接收第一信息之前，该方法还包括：该第一节点接收第四信息，该第四信息用于指示承载请求信息的资源，该请求信息用于请求该N个配置信息；该第一节点在该承载请求信息的资源上发送该请求信息。具体地，第一节点还可以接收用于指示承载请求信息的资源的第四信息，该请求信息用于请求该N个配置信息，使得该第一节点能够基于该第四信息指示的资源发送请求信息之后，能够通过第一信息获得该N个配置信息。

可选地，该请求信息包括第一正交序列，该第一正交序列对应于该本地信息。

可选地，该第一正交序列包含于P个正交序列，其中，该P个正交序列分别对应于不同的本地信息。换言之，不同第一节点均能够基于各自的本地信息在该P个正交序列中确定其中一个正交序列，并向第二节点发送该其中一个正交序列，使得该请求信息的接收方（例如第二节点）能够基于所接收的请求信息获知各个第一节点的本地信息。

可选地，在P大于1的情况下，不同序列可以是正交的关系，即P个序列可以表述为P个正交序列，以使该P个正交序列通过相同资源传输的情况下也能被接收方所识别。而在其它实现方式中，该P个序列也可以是通过其它方式被接收方所识别，例如通过频域资源、时域资源的错开。换言之，上述P个正交序列可以替换为P个序列，第一正交序列可以替换为第一序列。

此外，第二节点作为该请求信息的接收方的情况下，该第二节点能够基于接收的一个或多个正交序列确定各个第一节点的本地信息，可以将各个第一节点的本地信息作为N个配置信息以及M个配置信息以及K个配置信息至少一项的确定依据，使得该第二节点发送的各个配置信息能够与各个第一节点的本地信息相适配。

此外，由于不同的第一节点均能够通过该请求信息指示自身的本地参数规模，能够降低开销。

可选地，第一节点除了可以基于第四信息确定承载该第一正交序列的资源之外，该第一节点也可以通过其它方式（例如预配置的方式）确定承载该第一正交序列的资源，此处不做限定。

可选地，P的取值可以大于或等于N的取值。例如，第一节点可以通过配置或预配置的方式获知该P个正交序列分别对应的不同参数规模，并且，第二节点可以通过接收来自一个或多个第一节点的请求信息确

定存在各个第一节点的参数规模。相应的，该第二节点可以进一步基于所确定的各个第一节点的参数规模生成该N个配置信息。一般地，P个正交序列可以存在冗余设置（即P大于或等于N），以使得各个第一节点均能够通过请求信息指示各自的本地的参数规模。

在一种可能的实现方式中，在该第一节点接收第一信息之前，该方法还包括：该第一节点接收第五信息，该第五信息用于指示该N个模型参数分别对应的本地信息，和/或，该第五信息用于指示该第一信息是否包括该N个配置信息。具体地，第一节点还可以接收第五信息，以便于通过该第五信息获知该N个配置信息的相关信息，并且基于所获知的信息确定是否接收该N个配置信息，能够避免第一节点不必要的检测以降低开销。

作为一种实现示例，请求信息、第五信息以及第一信息可以通过不同的时域资源承载。如图4b所示，多个第一节点可以在请求信息所在的时域资源上发送请求信息，使得第二节点可以获得各个第一节点的本地的参数规模。此后，第二节点可以在第五信息中发送第一信息所指示的N个配置信息的相关信息，以便于第一节点基于该第五信息确定是否接收该N个配置信息，能够避免第一节点不必要的检测以降低开销。

在一种可能的实现方式中，在图3所示方法中，该方法还包括：该第一节点接收第二信息，该第二信息指示M个配置信息，该M个配置信息用于配置M个模型参数的传输资源，该M个模型参数的参数规模不同，M为大于1的整数；该第一节点基于该M个配置信息中的第二配置信息接收第二模型参数，其中，该第二模型参数的参数规模与本地的参数规模相同（或相近，例如该第二模型参数的参数规模与本地的参数规模的差低于阈值）。具体地，第一节点还可以接收指示M个配置信息的第二信息，并基于该M个配置信息的第二配置信息接收与本地的参数规模相同（或相近）的第二模型参数。换言之，不同的第一节点均可以通过第二信息接收与各自的本地信息的参数规模相同（或相近）的模型参数，并基于所接收的模型参数执行AI处理（例如本地训练）。

应理解，M个配置信息所配置的传输资源用于不同的第一节点接收M个模型参数中的其中一个模型参数，使得不同的第一节点后续均可以基于所接收的模型参数执行AI处理（例如本地训练）。相应的，该M个模型参数可以理解为多个第一节点的全局模型参数，即该M个模型参数可以替换为M个全局模型参数。

可选地，对于M个模型参数中的任一模型参数而言，可以包括梯度和/或权值等。例如，该梯度和/或权值用于初始化的神经网络模型（例如前文描述的第一神经网络模型）的进一步处理，例如神经网络模型的训练，神经网络模型的更新等。

类似地，M个模型参数的参数规模不同，可以替换为其它的描述，例如M个模型参数对应的神经网络层数不同，M个模型参数对应的神经元不同，M个模型参数对应的输入输出不同，M个模型参数的计算复杂度不同，或者，M个模型参数的计算能力要求不同，或者，M个模型参数的存储能力要求不同等。

在一种可能的实现方式中，该M个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的接收功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息。

具体地，第二信息指示的M个配置信息中的任意一个配置信息可以包括上述至少一项，以提升方案实现的灵活性。

可选地，第一节点在步骤S301中接收的第一信息指示的N个配置信息中的任意一个配置信息也可以包括以下至少一项，以提升方案实现的灵活性。该至少一项包括：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的接收功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息。

在一种可能的实现方式中，在该第一节点接收第二信息之前，该方法还包括：该第一节点接收第六信息，该第六信息用于指示该M个模型参数分别对应的本地信息，和/或，该第六信息用于指示该第三信息是

否包括该M个配置信息。具体地，第一节点还可以接收第六信息，以便于通过该第六信息获知该M个配置信息的相关信息，并且基于所获知的信息确定是否接收该M个配置信息，能够避免第一节点不必要的检测以降低开销。

在一种可能的实现方式中，在图3所示方法中，该方法还包括：该第一节点接收第三信息，该第三信息指示K个配置信息，该K个配置信息用于配置K个模型参数的传输资源，该K个模型参数的参数规模不同，K为大于1的整数；该第一节点基于该K个配置信息中的第三配置信息发送第三模型参数，其中，该第三模型参数的参数规模与本地的参数规模相同（或相近，例如该第三模型参数的参数规模与本地的参数规模的差低于阈值）。具体地，第一节点还可以接收指示K个配置信息的第三信息，并基于该K个配置信息的第三配置信息接收与本地的参数规模相同（或相近）的第三模型参数。换言之，不同的第一节点均可以通过第三信息发送与各自的本地信息的参数规模相同（或相近）的模型参数，使得该模型参数的接收方能够对不同的第一节点发送的模型参数执行AI处理（例如参数融合处理）。

应理解，K个配置信息所配置的传输资源用于不同的第一节点发送K个模型参数中的其中一个模型参数，使得该模型参数的接收方能够对不同的第一节点发送的模型参数执行融合操作。相应的，该K个模型参数可以理解多个第一节点的本地模型参数，即该K个模型参数可以替换为K个本地模型参数。

可选地，对于K个模型参数中的任一模型参数而言，可以包括梯度和/或权值等。例如，该梯度和/或权值用于神经网络模型的进一步融合处理，例如深度融合（Deep Fusion），加权平均融合（Weighted Averaging）等。

类似地，K个模型参数的参数规模不同，可以替换为其它描述，例如K个模型参数对应的神经网络层数不同，K个模型参数对应的神经元不同，K个模型参数对应的输入输出不同，K个模型参数的计算复杂度不同，或者，K个模型参数的计算能力要求不同，或者，K个模型参数的存储能力要求不同等。

在一种可能的实现方式中，该K个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的发射功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息；

指示该一个配置信息的传输资源在该K个配置信息配置的传输资源中的索引的指示信息。

应理解，在该任意的一个配置信息包括指示该一个配置信息的传输资源在该K个配置信息配置的传输资源中的索引的指示信息的情况下，可以理解为第二节点为该K个配置信息配置了传输资源集合（包括时域资源、频域资源、空域资源中的至少一项），使得第一节点能够基于该指示信息的索引在该传输资源集合中确定相应的传输资源。

具体地，第三信息指示的K个配置信息中的任意一个配置信息可以包括上述至少一项，以提升方案实现的灵活性。

在一种可能的实现方式中，在该第一节点接收第三信息之前，该方法还包括：该第一节点接收第七信息，该第七信息用于指示该K个模型参数分别对应的本地信息，和/或，该第七信息用于指示该第二信息是否包括该K个配置信息。具体地，第一节点还可以接收第七信息，以便于通过该第七信息获知该K个配置信息的相关信息，并且基于所获知的信息确定是否接收该K个配置信息，能够避免第一节点不必要的检测以降低开销。

可选地，N/M/K这三者的取值可以互不相同，也可以是至少两个相同，也可以是三个都相同，此处不做限定。

作为一种实现示例，以M个配置信息所配置的资源用于传输M个全局模型参数，且K个配置信息所配置的资源用于传输N个本地模型参数为例。如图4c所示，在M和K的取值均为4的情况下，在时域上，Cat0的第一节点、Cat1的第一节点、Cat2的第一节点、Cat3的第一节点分别可以在相应的时域资源上接收M个全局模型参数中的其中一个全局模型参数，以便于各个第一节点基于该全局模型参数在本地执行AI处理（例如AI训练处理）。此后，在时域上，Cat0的第一节点、Cat1的第一节点、Cat2的第一节点、Cat3的第

一节点分别可以在相应的时域资源上发送K个本地模型参数中的其中一个本地模型参数，以便于第二节点基于接收的各个本地模型参数在执行AI处理（例如AI融合处理）。

基于图3所示技术方案，第一节点在步骤S301中接收的第一信息指示N个配置信息，该N个配置信息用于配置N个模型参数的资源，此后，该第一节点在步骤S302中基于该N个配置信息中的第一配置信息所配置的资源接收第一模型参数，并且，该第一节点可以基于该第一模型参数确定与本地信息对应的第一神经网络模型。换言之，不同的第一节点均可以基于该第一信息获得与各自的本地信息对应的神经网络模型，后续均可以基于本地获得的神经网络模型参与AI学习。相比于不同节点基于随机化的神经网络模型参与AI学习的实现方式，在上述技术方案中，不同节点可以基于与各自的本地信息对应的神经网络模型参与AI学习。从而，在通信系统中的通信节点作为参与AI学习的节点的情况下，能够使得通信节点的算力能够应用于学习系统中的AI任务，并能够提升不同节点的实现灵活性。

此外，不同节点的本地信息一般是不同的，而在上述技术方案中，第一节点能够基于第一信息确定与本地信息对应的神经网络模型，为此，不同的第一节点均能够基于与本地信息适配的神经网络模型参与AI学习，以提升神经网络模型的处理效率。

请参阅图5，为本申请提供的通信方法的另一个实现示意图，该方法包括如下步骤。

需要说明的是，图5中以第一节点和第二节点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，图5及对应的实现方式中S501的执行主体是第一节点，执行主体也可以是支持该第一节点实现该方法的芯片、芯片系统、或处理器，还可以是能够实现全部或部分第一节点功能的逻辑模块或软件。图5及对应的实现方式中S501-S502中的第二节点也可以替换为支持该第二节点实现该方法的芯片、芯片系统、或处理器，还可以替换为能够实现全部或部分第二节点功能的逻辑模块或软件。

S501.第二节点发送第二信息，相应的，第一节点接收该第二信息。其中，该第二信息用于指示M个配置信息，该M个配置信息用于配置M个模型参数的传输资源，该M个模型参数的参数规模不同，M为大于1的整数。

S502.第一节点基于M个配置信息中的第二配置信息接收第二模型参数。其中，该第二模型参数的参数规模与本地的参数规模相同（或相近，例如该第二模型参数的参数规模与本地的参数规模的差低于阈值）。

基于图5所示技术方案，第一节点在步骤S501中接收的第二信息指示M个配置信息，该M个配置信息用于配置M个模型参数的资源，此后，在步骤S502中，该第一节点基于该M个配置信息中的第二配置信息所配置的资源接收第二模型参数，并且，该第二模型参数与本地的参数规模相同（或相近）。换言之，不同的第一节点均可以基于该第一信息获得与本地的参数规模相同（或相近）的模型参数，后续均可以基于与本地的参数规模相同（或相近）的模型参数参与AI学习。相比于不同节点基于随机化的模型参数参与AI学习的实现方式，在上述技术方案中，不同节点可以基于与本地的参数规模相同（或相近）的模型参数参与AI学习。从而，在通信系统中的通信节点作为参与AI学习的节点的情况下，能够使得通信节点的算力能够应用于学习系统中的AI任务，并能够提升不同节点的实现灵活性。

此外，不同节点的本地参数规模可能是不同的，在上述技术方案中，第一节点能够基于与本地的参数规模相同（或相近）的模型参数参与AI学习，为此，不同的第一节点均能够基于与本地信息适配的模型参数参与AI学习，以提升神经网络模型的处理效率。

需要说明的是，图5所示技术方案可以参考前述图3及相关实施例的描述，并实现相应的技术效果，例如本地信息、模型参数、M个配置信息、第三信息、K个配置信息、第六信息、第七信息等。

请参阅图6，为本申请提供的通信方法的另一个实现示意图，该方法包括如下步骤。

需要说明的是，图6中以第一节点和第二节点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，图6及对应的实现方式中S601的执行主体是第一节点，执行主体也可以是支持该第一节点实现该方法的芯片、芯片系统、或处理器，还可以是能够实现全部或部分第一节点功能的逻辑模块或软件。图6及对应的实现方式中S601-S602中的第二节点也可以替换为支持该第二节点

实现该方法的芯片、芯片系统、或处理器，还可以替换为能实现全部或部分第二节点功能的逻辑模块或软件。

S601.第二节点发送第三信息，相应的，第一节点接收该第三信息。其中，该第三信息指示K个配置信息，该K个配置信息用于配置K个模型参数的传输资源，该K个模型参数的参数规模不同，K为大于1的整数。

S602.第一节点基于该K个配置信息中的第三配置信息发送第三模型参数，其中，该第三模型参数的参数规模与本地的参数规模相同（或相近，例如该第三模型参数的参数规模与本地的参数规模的差低于阈值）。

基于图6所示技术方案，第一节点在步骤S601中接收的第三信息指示K个配置信息，该K个配置信息用于配置K个模型参数的资源，此后，在步骤S602中，该第一节点基于该K个配置信息中的第三配置信息所配置的资源发送第三模型参数，并且，该第三模型参数与本地的参数规模相同（或相近）。换言之，不同的第一节点均可以基于该第一信息获得与本地的参数规模相同（或相近）的模型参数，后续均可以基于与本地的参数规模相同（或相近）的模型参数参与AI学习。相比于不同节点基于随机化的模型参数参与AI学习的实现方式，在上述技术方案中，不同节点可以基于与本地的参数规模相同（或相近）的模型参数参与AI学习。从而，在通信系统中的通信节点作为参与AI学习的节点的情况下，能够使得通信节点的算力能够应用于学习系统中的AI任务，并能够提升不同节点的实现灵活性。

此外，不同节点的本地参数规模可能是不同的，在上述技术方案中，第一节点能够基于与本地的参数规模相同（或相近）的模型参数参与AI学习，为此，不同的第一节点均能够基于与本地信息适配的模型参数参与AI学习，以提升神经网络模型的处理效率。

需要说明的是，图6所示技术方案可以参考前述图3及相关实施例的描述，并实现相应的技术效果，例如本地信息、模型参数、第三信息、K个配置信息、第七信息等。

作为前文图3至图6所示技术方案的一种实现示例，如图7所示，第二节点和第一节点之间可以通过图7所示时域的流向执行相关步骤。

一种实现示例中，第二节点可以通过调度（Scheduling）的资源授权（Grant）方式向第一节点发送图3所示技术方案涉及的第一信息，使得第一节点基于第一信息指示的N个配置信息获得模型参数，并基于该模型参数执行初始化过程以得到第一神经网络模型，后续可以基于该第一神经网络模型执行AI处理过程（例如第一节点的AI训练处理，第二节点的AI融合处理等）。

可选地，在第二节点调度第一信息之前，第二节点还可以向第一节点发送第四信息，使得第一节点基于该第四信息配置的资源发送请求信息，该请求信息可以用于触发第二节点调度N个配置信息所指示的资源。

另一种实现示例中，第二节点可以通过调度（Scheduling）的资源授权（Grant）方式向第一节点发送图3/图5/图6所示技术方案涉及的第二信息，使得第一节点基于第二信息指示的M个配置信息接收全局模型参数，并基于该全局模型参数执行AI处理过程（例如第一节点的AI训练处理，第二节点的AI融合处理等）。

另一种实现示例中，第二节点可以通过调度（Scheduling）的资源授权（Grant）方式向第一节点发送图3/图5/图6所示技术方案涉及的第三信息，使得第一节点基于第三信息指示的K个配置信息发送本地模型参数，并基于该全局模型参数执行AI处理过程（例如第一节点的AI训练处理，第二节点的AI融合处理等）。

可选地，在AI处理过程中，第一节点可以向第二节点反馈（Feedback）本地的AI处理结果，相应的，第二节点也可以向第一节点调整（Adjustment）AI处理的相关信息（例如模型梯度/模型权值等）。

可选地，在AI处理过程之后，第二节点还可以发送指示（indicator），该指示用于指示AI处理终止（terminated）。换言之，对于第一节点而言，第一节点可以通过事件检测（Event monitoring）的方式，检测到该指示之后可以确定AI处理终止。

请参阅图8，本申请实施例提供了一种通信装置800，该通信装置800可以实现上述方法实施例中第一节点或第二节点（该第一节点或第二节点为终端设备或网络设备）的功能，因此也能实现上述方法实施例

所具备的有益效果。在本申请实施例中，该通信装置800可以是第一节点（或第二节点），也可以是第一节点（或第二节点）内部的集成电路或者元件等，例如芯片。下文实施例以该通信装置800为第一节点（或第二节点）为例进行说明。

需要说明的是，收发单元802可以包括发送单元和接收单元，分别用于执行发送和接收。

5 一种可能的实现方式中，当该装置800为用于执行前述任一实施例中第一节点所执行的方法时，该装置800包括处理单元801和收发单元802；该收发单元802用于接收第一信息，该第一信息指示N个配置信息，该N个配置信息用于配置N个模型参数的资源，N为正整数；该处理单元801用于确定该N个配置信息中的第一配置信息；该收发单元802还用于基于该第一配置信息接收第一模型参数，该第一模型参数用于确定与本地信息对应的第一神经网络模型。

10 在一种可能的实现方式中，该本地信息包括本地资源开销水平，本地AI能力信息、本地的参数规模中的至少一项。

在一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

15 在一种可能的实现方式中，N为大于1的整数，该N个模型参数的参数规模不同，该处理单元801还用于基于该本地信息在该N个配置信息中确定该第一配置信息。

在一种可能的实现方式中，N的取值为1，其中，该第一模型参数用于配置至少两个第一节点共用的神经网络模型；该处理单元801还用于基于该本地信息将该至少两个第一节点共用的神经网络模型进行知识蒸馏处理，得到该第一神经网络模型。

20 在一种可能的实现方式中，该收发单元802还用于接收第二信息，该第二信息指示M个配置信息，该M个配置信息用于配置M个模型参数的传输资源，该M个模型参数的参数规模不同，M为大于1的整数；该处理单元801还用于确定该M个配置信息中的第二配置信息；该收发单元802还用于基于该第二配置信息接收第二模型参数，其中，该第二模型参数的参数规模与本地的参数规模相同（或相近）。

在一种可能的实现方式中，该M个配置信息中的任意一个配置信息包括以下至少一项：

25 指示该一个配置信息配置的模型参数的参数规模的指示信息；
指示该一个配置信息配置的传输资源的接收功率的指示信息；
指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；
指示该一个配置信息配置的传输资源的频域资源的指示信息；
指示该一个配置信息配置的传输资源的时域资源的指示信息。

30 在一种可能的实现方式中，该收发单元802还用于接收第三信息，该第三信息指示K个配置信息，该K个配置信息用于配置K个模型参数的传输资源，该K个模型参数的参数规模不同，K为大于1的整数；该处理单元801还用于确定该K个配置信息中的第三配置信息；该收发单元802还用于基于该第三配置信息发送第三模型参数，其中，该第三模型参数的参数规模与本地的参数规模相同（或相近）。

在一种可能的实现方式中，该K个配置信息中的任意一个配置信息包括以下至少一项：

35 指示该一个配置信息配置的模型参数的参数规模的指示信息；
指示该一个配置信息配置的传输资源的发射功率的指示信息；
指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；
指示该一个配置信息配置的传输资源的频域资源的指示信息；
指示该一个配置信息配置的传输资源的时域资源的指示信息；
指示该一个配置信息的传输资源在该K个配置信息配置的传输资源中的索引的指示信息。

40 在一种可能的实现方式中，该收发单元802还用于接收第四信息，该第四信息用于指示承载请求信息的资源，该请求信息用于请求该N个配置信息；该收发单元802还用于在该承载请求信息的资源上发送该请求信息。

在一种可能的实现方式中，该请求信息包括第一正交序列，该第一正交序列对应于该本地信息。

45 在一种可能的实现方式中，该收发单元802还用于接收第五信息，该第五信息用于指示该N个模型参数分别对应的本地信息，和/或，该第五信息用于指示该第一信息是否包括该N个配置信息。

一种可能的实现方式中，当该装置800为用于执行前述任一实施例第二节节点所执行的方法时，该装置800包括处理单元801和收发单元802；该处理单元801用于确定第一信息，该第一信息指示N个配置信息，该N个配置信息用于配置N个模型参数的资源，N为正整数；其中，该N个模型参数中的一个模型参数用于确定与第一节节点的本地信息对应的神经网络模型；该收发单元802用于发送该第一信息。

5 在一种可能的实现方式中，该本地信息包括本地资源开销水平，本地AI能力信息、本地的参数规模中的至少一项。

在一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

在一种可能的实现方式中，N为大于1的整数，该N个模型参数的参数规模不同。

10 在一种可能的实现方式中，N的取值为1。

在一种可能的实现方式中，该收发单元802还用于发送第二信息，该第二信息指示M个配置信息，该M个配置信息用于配置M个模型参数的传输资源，该M个模型参数的参数规模不同，M为大于1的整数；该收发单元802还用于基于该M个配置信息发送该M个模型参数的部分或全部。

在一种可能的实现方式中，该M个配置信息中的任意一个配置信息包括以下至少一项：

15 指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的接收功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息。

20 在一种可能的实现方式中，该收发单元802还用于发送第三信息，该第三信息指示K个配置信息，该K个配置信息用于配置K个模型参数的传输资源，该K个模型参数的参数规模不同，K为大于1的整数；该收发单元802还用于基于该K个配置信息接收该K个模型参数的部分或全部。

在一种可能的实现方式中，该K个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

25 指示该一个配置信息配置的传输资源的发射功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息；

指示该一个配置信息的传输资源在该K个配置信息配置的传输资源中的索引的指示信息。

30 在一种可能的实现方式中，该收发单元802还用于发送第四信息，该第四信息用于指示承载请求信息的资源，该请求信息用于请求该N个配置信息；该收发单元802还用于在该承载请求信息的资源上接收该请求信息。

在一种可能的实现方式中，该请求信息包括第一正交序列，该第一正交序列对应于该本地信息。

35 在一种可能的实现方式中，该收发单元802还用于发送第五信息，该第五信息用于指示该N个模型参数分别对应的本地信息，和/或，该第五信息用于指示该第一信息是否包括该N个配置信息。

一种可能的实现方式中，当该装置800为用于执行前述任一实施例第一节节点所执行的方法时，该装置800包括处理单元801和收发单元802；该收发单元802用于接收第二信息，该第二信息指示M个配置信息，该M个配置信息用于配置M个模型参数的传输资源，该M个模型参数的参数规模不同，M为大于1的整数；该处理单元801用于确定M个配置信息中的第二配置信息；该收发单元802还用于基于该第二配置信息接收第二模型参数，其中，该第二模型参数的参数规模与本地的参数规模相同（或相近）。

40 在一种可能的实现方式中，该本地信息包括本地资源开销水平，本地AI能力信息、本地的参数规模中的至少一项。

在一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

45 在一种可能的实现方式中，该M个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的接收功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息。

- 5 在一种可能的实现方式中，该收发单元802还用于接收第三信息，该第三信息指示K个配置信息，该K个配置信息用于配置K个模型参数的传输资源，该K个模型参数的参数规模不同，K为大于1的整数；该收发单元802还用于基于该K个配置信息中的第三配置信息发送第三模型参数，其中，该第三模型参数的参数规模与本地的参数规模相同（或相近）。

在一种可能的实现方式中，该K个配置信息中的任意一个配置信息包括以下至少一项：

- 10 指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的发射功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息；

- 15 指示该一个配置信息的传输资源在该K个配置信息配置的传输资源中的索引的指示信息。

一种可能的实现方式中，当该装置800为用于执行前述任一实施例第二节所执行的方法时，该装置800包括处理单元801和收发单元802；该处理单元801用于确定第二信息，该第二信息指示M个配置信息，该M个配置信息用于配置M个模型参数的传输资源，该M个模型参数的参数规模不同，M为大于1的整数；该收发单元802用于发送第二信息。

- 20 在一种可能的实现方式中，该本地信息包括本地资源开销水平，本地AI能力信息、本地的参数规模中的至少一项。

在一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

在一种可能的实现方式中，该M个配置信息中的任意一个配置信息包括以下至少一项：

- 25 指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的接收功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息。

- 30 在一种可能的实现方式中，该处理单元801还用于确定第三信息，该第三信息指示K个配置信息，该K个配置信息用于配置K个模型参数的传输资源，该K个模型参数的参数规模不同，K为大于1的整数；

在一种可能的实现方式中，该K个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

指示该一个配置信息配置的传输资源的发射功率的指示信息；

- 35 指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息；

指示该一个配置信息的传输资源在该K个配置信息配置的传输资源中的索引的指示信息。

一种可能的实现方式中，当该装置800为用于执行前述任一实施例第一节所执行的方法时，该装置800包括处理单元801和收发单元802；该收发单元802用于接收第三信息，该第三信息指示K个配置信息，该K个配置信息用于配置K个模型参数的传输资源，该K个模型参数的参数规模不同，K为大于1的整数；该处理单元801用于确定该K个配置信息中的第三配置信息；该收发单元802还用于基于该第三配置信息发送第三模型参数，其中，该第三模型参数的参数规模与本地的参数规模相同（或相近）。

- 40 在一种可能的实现方式中，该本地信息包括本地资源开销水平，本地AI能力信息、本地的参数规模中的至少一项。

45

在一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

在一种可能的实现方式中，该K个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

5 指示该一个配置信息配置的传输资源的发射功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息；

指示该一个配置信息的传输资源在该K个配置信息配置的传输资源中的索引的指示信息。

10 一种可能的实现方式中，当该装置800为用于执行前述任一实施例第二节所执行的方法时，该装置800包括处理单元801和收发单元802；该处理单元801用于确定第三信息，该第三信息指示K个配置信息，该K个配置信息用于配置K个模型参数的传输资源，该K个模型参数的参数规模不同，K为大于1的整数；该收发单元802用于发送第三信息。

15 在一种可能的实现方式中，该本地信息包括本地资源开销水平，本地AI能力信息、本地的参数规模中的至少一项。

在一种可能的实现方式中，该模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

在一种可能的实现方式中，该K个配置信息中的任意一个配置信息包括以下至少一项：

指示该一个配置信息配置的模型参数的参数规模的指示信息；

20 指示该一个配置信息配置的传输资源的发射功率的指示信息；

指示该一个配置信息配置的模型参数的参数类型的指示信息，该参数类型包括梯度和/或权值；

指示该一个配置信息配置的传输资源的频域资源的指示信息；

指示该一个配置信息配置的传输资源的时域资源的指示信息；

指示该一个配置信息的传输资源在该K个配置信息配置的传输资源中的索引的指示信息。

25 需要说明的是，上述通信装置800的单元的信息执行过程等内容，具体可参见本申请前述所示的方法实施例中的叙述，此处不再赘述。

请参阅图9，为本申请提供的通信装置900的另一种示意性结构图，通信装置900至少包括输入输出接口902。其中，通信装置900可以为芯片或集成电路。

30 可选的，该通信装置还包括逻辑电路901。

其中，图8所示收发单元802可以为通信接口，该通信接口可以是图9中的输入输出接口902，该输入输出接口902可以包括输入接口和输出接口。或者，该通信接口也可以是收发电路，该收发电路可以包括输入接口电路和输出接口电路。

35 可选的，输入输出接口902用于接收第一信息，该第一信息指示N个配置信息，该N个配置信息用于配置N个模型参数的资源，N为正整数；逻辑电路901用于确定该N个配置信息中的第一配置信息；该输入输出接口902还用于基于该第一配置信息接收第一模型参数，该第一模型参数用于确定与本地信息对应的第一神经网络模型。

40 可选地，逻辑电路901用于确定第一信息，该第一信息指示N个配置信息，该N个配置信息用于配置N个模型参数的资源，N为正整数；其中，该N个模型参数中的一个模型参数用于确定与第一节节点的本地信息对应的神经网络模型；输入输出接口902用于发送该第一信息。

可选地，输入输出接口902用于接收第二信息，该第二信息指示M个配置信息，该M个配置信息用于配置M个模型参数的传输资源，该M个模型参数的参数规模不同，M为大于1的整数；该逻辑电路901用于确定M个配置信息中的第二配置信息；该输入输出接口902还用于基于该第二配置信息接收第二模型参数，其中，该第二模型参数的参数规模与本地的参数规模相同（或相近）。

可选地，逻辑电路901用于确定第二信息，该第二信息指示M个配置信息，该M个配置信息用于配置M个模型参数的传输资源，该M个模型参数的参数规模不同，M为大于1的整数；该输入输出接口902用于发送第二信息。

5 可选地，输入输出接口902用于接收第三信息，该第三信息指示K个配置信息，该K个配置信息用于配置K个模型参数的传输资源，该K个模型参数的参数规模不同，K为大于1的整数；逻辑电路901用于确定该K个配置信息中的第三配置信息；该输入输出接口902还用于基于该第三配置信息发送第三模型参数，其中，该第三模型参数的参数规模与本地的参数规模相同（或相近）。

10 可选地，逻辑电路901用于确定第三信息，该第三信息指示K个配置信息，该K个配置信息用于配置K个模型参数的传输资源，该K个模型参数的参数规模不同，K为大于1的整数；输入输出接口902用于发送第三信息。

其中，逻辑电路901和输入输出接口902还可以执行任一实施例中第一节点或第二节点执行的其他步骤并实现对应的有益效果，此处不再赘述。

在一种可能的实现方式中，图8所示处理单元801可以为图9中的逻辑电路901。

15 可选的，逻辑电路901可以是一个处理装置，处理装置的功能可以部分或全部通过软件实现。其中，处理装置的功能可以部分或全部通过软件实现。

可选的，处理装置可以包括存储器和处理器，其中，存储器用于存储计算机程序，处理器读取并执行存储器中存储的计算机程序，以执行任意一个方法实施例中的相应处理和/或步骤。

20 可选地，处理装置可以仅包括处理器。用于存储计算机程序的存储器位于处理装置之外，处理器通过电路/电线与存储器连接，以读取并执行存储器中存储的计算机程序。其中，存储器和处理器可以集成在一起，或者也可以是物理上互相独立的。

可选地，该处理装置可以是一个或多个芯片，或一个或多个集成电路。例如，处理装置可以是一个或多个现场可编程门阵列（field-programmable gate array, FPGA）、专用集成芯片（application specific integrated circuit, ASIC）、系统芯片（system on chip, SoC）、中央处理器（central processor unit, CPU）、网络处理器（network processor, NP）、数字信号处理电路（digital signal processor, DSP）、微控制器（micro controller unit, MCU）、可编程控制器（programmable logic device, PLD）或其它集成芯片，或者上述芯片或者处理器的任意组合等。

30 请参阅图10，为本申请的实施例提供的上述实施例中所涉及的通信装置1000，该通信装置1000具体可以为上述实施例中的作为终端设备的通信装置，图10所示示例为终端设备通过终端设备（或者终端设备中的部件）实现。

其中，该通信装置1000的一种可能的逻辑结构示意图，该通信装置1000可以包括但不限于至少一个处理器1001以及通信端口1002。

35 其中，图8所示收发单元802可以为通信接口，该通信接口可以是图10中的通信端口1002，该通信端口1002可以包括输入接口和输出接口。或者，该通信端口1002也可以是收发电路，该收发电路可以包括输入接口电路和输出接口电路。

进一步可选的，该装置还可以包括存储器1003、总线1004中的至少一个，在本申请的实施例中，该至少一个处理器1001用于对通信装置1000的动作进行控制处理。

40 此外，处理器1001可以是中央处理器单元，通用处理器，数字信号处理器，专用集成电路，现场可编程门阵列或者其他可编程逻辑器件、晶体管逻辑器件、硬件部件或者其任意组合。其可以实现或执行结合本申请公开内容所描述的各种示例性的逻辑方框，模块和电路。该处理器也可以是实现计算功能的组合，例如包含一个或多个微处理器组合，数字信号处理器和微处理器的组合等等。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，上述描述的系统，装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

45 需要说明的是，图10所示通信装置1000具体可以用于实现前述方法实施例中终端设备所实现的步骤，并实现终端设备对应的技术效果，图10所示通信装置的具体实现方式，均可以参考前述方法实施例中的叙述，此处不再一一赘述。

请参阅图11，为本申请的实施例提供的上述实施例中所涉及的通信装置1100的结构示意图，该通信装置1100具体可以为上述实施例中的作为网络设备的通信装置，图11所示示例为网络设备通过网络设备（或者网络设备中的部件）实现，其中，该通信装置的结构可以参考图11所示的结构。

5 通信装置1100包括至少一个处理器1111以及至少一个网络接口1114。进一步可选的，该通信装置还包括至少一个存储器1112、至少一个收发器1113和一个或多个天线1115。处理器1111、存储器1112、收发器1113和网络接口1114相连，例如通过总线相连，在本申请实施例中，该连接可包括各类接口、传输线或总线等，本实施例对此不做限定。天线1115与收发器1113相连。网络接口1114用于使得通信装置通过通信链路，与其它通信设备通信。例如网络接口1114可以包括通信装置与核心网设备之间的网络接口，例如S1接口，网络接口可以包括通信装置和其他通信装置（例如其他网络设备或者核心网设备）之间的网络接口，
10 例如X2或者Xn接口。

其中，图8所示收发单元802可以为通信接口，该通信接口可以是图11中的网络接口1114，该网络接口1114可以包括输入接口和输出接口。或者，该网络接口1114也可以是收发电路，该收发电路可以包括输入接口电路和输出接口电路。

15 处理器1111主要用于对通信协议以及通信数据进行处理，以及对整个通信装置进行控制，执行软件程序，处理软件程序的数据，例如用于支持通信装置执行实施例中所描述的动作。通信装置可以包括基带处理器和中央处理器，基带处理器主要用于对通信协议以及通信数据进行处理，中央处理器主要用于对整个终端设备进行控制，执行软件程序，处理软件程序的数据。图11中的处理器1111可以集成基带处理器和中央处理器的功能，本领域技术人员可以理解，基带处理器和中央处理器也可以是各自独立的处理器，通过
20 总线等技术互联。本领域技术人员可以理解，终端设备可以包括多个基带处理器以适应不同的网络制式，终端设备可以包括多个中央处理器以增强其处理能力，终端设备的各个部件可以通过各种总线连接。该基带处理器也可以表述为基带处理电路或者基带处理芯片。该中央处理器也可以表述为中央处理电路或者中央处理芯片。对通信协议以及通信数据进行处理的功能可以内置在处理器中，也可以以软件程序的形式存储在存储器中，由处理器执行软件程序以实现基带处理功能。

25 存储器主要用于存储软件程序和数据。存储器1112可以是独立存在，与处理器1111相连。可选的，存储器1112可以和处理器1111集成在一起，例如集成在一个芯片之内。其中，存储器1112能够存储执行本申请实施例的技术方案的程序代码，并由处理器1111来控制执行，被执行的各类计算机程序代码也可被视为是处理器1111的驱动程序。

图11仅示出了一个存储器和一个处理器。在实际的终端设备中，可以存在多个处理器和多个存储器。
30 存储器也可以称为存储介质或者存储设备等。存储器可以为与处理器处于同一芯片上的存储元件，即片内存储元件，或者为独立的存储元件，本申请实施例对此不做限定。

收发器1113可以用于支持通信装置与终端之间射频信号的接收或者发送，收发器1113可以与天线1115相连。收发器1113包括发射机Tx和接收机Rx。具体地，一个或多个天线1115可以接收射频信号，该收发器1113的接收机Rx用于从天线接收该射频信号，并将射频信号转换为数字基带信号或数字中频信号，并将该
35 数字基带信号或数字中频信号提供给该处理器1111，以便处理器1111对该数字基带信号或数字中频信号做进一步的处理，例如解调处理和译码处理。此外，收发器1113中的发射机Tx还用于从处理器1111接收经过调制的数字基带信号或数字中频信号，并将该经过调制的数字基带信号或数字中频信号转换为射频信号，并通过一个或多个天线1115发送该射频信号。具体地，接收机Rx可以选择性地对射频信号进行一级或多级下混频处理和模数转换处理以得到数字基带信号或数字中频信号，该下混频处理和模数转换处理的先后顺
40 序是可调整的。发射机Tx可以选择性地对经过调制的数字基带信号或数字中频信号时进行一级或多级上混频处理和数模转换处理以得到射频信号，该上混频处理和数模转换处理的先后顺序是可调整的。数字基带信号和数字中频信号可以统称为数字信号。

收发器1113也可以称为收发单元、收发机、收发装置等。可选的，可以将收发单元中用于实现接收功能的器件视为接收单元，将收发单元中用于实现发送功能的器件视为发送单元，即收发单元包括接收单元和发送单元，接收单元也可以称为接收机、输入口、接收电路等，发送单元可以称为发射机、发射器或者
45 发射电路等。

需要说明的是，图11所示通信装置1100具体可以用于实现前述方法实施例中网络设备所实现的步骤，并实现网络设备对应的技术效果，图11所示通信装置1100的具体实现方式，均可以参考前述方法实施例中的叙述，此处不再一一赘述。

请参阅图12，为本申请的实施例提供的上述实施例中所涉及的通信装置的结构示意图。

5 可以理解的是，通信装置120包括例如模块、单元、元件、电路、或接口等，以适当地配置在一起以执行本申请提供的技术方案。所述通信装置120可以是前文描述的RAN节点、终端、核心网设备或者其他网络设备，也可以是这些设备中的部件（例如芯片），用以实现下述方法实施例中描述的方法。通信装置120包括一个或多个处理器121。所述处理器121可以是通用处理器或者专用处理器等。例如可以是基带处理器、或中央处理器。基带处理器可以用于对通信协议以及通信数据进行处理，中央处理器可以用于对通信装置（如，RAN节点、终端、或芯片等）进行控制，执行软件程序，处理软件程序的数据。

10 可选的，在一种设计中，处理器121可以包括程序123（有时也可以称为代码或指令），所述程序123可以在所述处理器121上被运行，使得所述通信装置120执行下述实施例中描述的方法。在又一种可能的设计中，通信装置120包括电路（图12未示出）。

15 可选的，所述通信装置120中可以包括一个或多个存储器122，其上存有程序124（有时也可以称为代码或指令），所述程序124可在所述处理器121上被运行，使得所述通信装置120执行上述方法实施例中描述的方法。

可选的，所述处理器121和/或存储器122中可以包括AI模块127,128，所述AI模块用于实现AI相关的功能。所述AI模块可以通过软件，硬件，或软硬结合的方式实现。例如，AI模块可以包括无线智能控制（radio intelligence control，RIC）模块。例如AI模块可以是近实时RIC或者非实时RIC。

20 可选的，所述处理器121和/或存储器122中还可以存储有数据。所述处理器和存储器可以单独设置，也可以集成在一起。

可选的，所述通信装置120还可以包括收发器125和/或天线126。所述处理器121有时也可以称为处理单元，对通信装置（例如RAN节点或终端）进行控制。所述收发器125有时也可以称为收发单元、收发机、收发电路、或者收发器等，用于通过天线126实现通信装置的收发功能。

25 其中，图8所示收发单元802可以为通信接口，该通信接口可以是图12中的收发器125，该收发器125可以包括输入接口和输出接口。或者，该收发器125也可以是收发电路，该收发电路可以包括输入接口电路和输出接口电路。

30 本申请实施例还提供一种计算机可读存储介质，该存储介质用于存储一个或多个计算机执行指令，当计算机执行指令被处理器执行时，该处理器执行如前述实施例中第一节点或第二节点可能的实现方式所述的方法。

本申请实施例还提供一种计算机程序产品（或称计算机程序），当计算机程序产品被该处理器执行时，该处理器执行上述第一节点或第二节点可能实现方式的方法。

35 本申请实施例还提供了一种芯片系统，该芯片系统包括至少一个处理器，用于支持通信装置实现上述通信装置可能的实现方式中所涉及的功能。可选的，所述芯片系统还包括接口电路，所述接口电路为所述至少一个处理器提供程序指令和/或数据。在一种可能的设计中，该芯片系统还可以包括存储器，存储器，用于保存该通信装置必要的程序指令和数据。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，其中，该通信装置具体可以为前述方法实施例中第一节点或第二节点。

本申请实施例还提供了一种通信系统，该网络系统架构包括上述任一实施例中的第一节点和第二节点，该第一节点可以为终端设备或网络设备，该第二节点也可以为终端设备或网络设备。

40 在本申请所提供的几个实施例中，应该理解到，所揭露的系统，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系统，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或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。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。所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。基于这样的理解，本申请的技术方案本质上或者说做出贡献的部分或者该技术方案的全部或部分可以以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中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等）执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U盘、移动硬盘、只读存储器（ROM，Read-Only Memory）、随机存取存储器（RAM，Random Access Memory）、磁碟或者光盘等各种可以存储程序代码的介质。

权 利 要 求

1. 一种通信方法，其特征在于，包括：

接收第一信息，所述第一信息指示 N 个配置信息，所述 N 个配置信息用于配置 N 个模型参数的资源，N 为正整数；

5 基于所述 N 个配置信息中的第一配置信息接收第一模型参数，所述第一模型参数用于确定与本地信息对应的第一神经网络模型。

2. 根据权利要求 1 所述的方法，其特征在于，所述本地信息包括本地资源开销水平，本地 AI 能力信息、本地的参数规模中的至少一项。

10 3. 根据权利要求 1 或 2 所述的方法，其特征在于，所述模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

15 4. 根据权利要求 1 至 3 任一项所述的方法，其特征在于，N 为大于 1 的整数，所述 N 个模型参数的参数规模不同，所述方法还包括：

基于所述本地信息在所述 N 个配置信息中确定所述第一配置信息。

5. 根据权利要求 1 至 3 任一项所述的方法，其特征在于，N 的取值为 1，其中，所述第一模型参数用于配置至少两个第一节点共用的神经网络模型；所述方法还包括：

20 基于所述本地信息将所述至少两个第一节点共用的神经网络模型进行知识蒸馏处理，得到所述第一神经网络模型。

6. 根据权利要求 1 至 5 任一项所述的方法，其特征在于，所述方法还包括：

25 接收第二信息，所述第二信息指示 M 个配置信息，所述 M 个配置信息用于配置 M 个模型参数的传输资源，所述 M 个模型参数的参数规模不同，M 为大于 1 的整数；

基于所述 M 个配置信息中的第二配置信息接收第二模型参数，其中，所述第二模型参数的参数规模与本地的参数规模相同。

30 7. 根据权利要求 6 所述的方法，其特征在于，所述 M 个配置信息中的任意一个配置信息包括以下至少一项：

- 指示所述一个配置信息配置的模型参数的参数规模的指示信息；
- 指示所述一个配置信息配置的传输资源的接收功率的指示信息；
- 指示所述一个配置信息配置的模型参数的参数类型的指示信息，所述参数类型包括梯度和/或权值；
- 指示所述一个配置信息配置的传输资源的频域资源的指示信息；
- 35 指示所述一个配置信息配置的传输资源的时域资源的指示信息。

8. 根据权利要求 1 至 7 任一项所述的方法，其特征在于，所述方法还包括：

接收第三信息，所述第三信息指示 K 个配置信息，所述 K 个配置信息用于配置 K 个模型参数的传输资源，所述 K 个模型参数的参数规模不同，K 为大于 1 的整数；

40 基于所述 K 个配置信息中的第三配置信息发送第三模型参数，其中，所述第三模型参数的参数规模与本地的参数规模相同。

9. 根据权利要求 8 所述的方法，其特征在于，所述 K 个配置信息中的任意一个配置信息包括以下至少一项：

- 45 指示所述一个配置信息配置的模型参数的参数规模的指示信息；
- 指示所述一个配置信息配置的传输资源的发射功率的指示信息；

指示所述一个配置信息配置的模型参数的参数类型的指示信息，所述参数类型包括梯度和/或权值；
指示所述一个配置信息配置的传输资源的频域资源的指示信息；
指示所述一个配置信息配置的传输资源的时域资源的指示信息；
指示所述一个配置信息的传输资源在所述 K 个配置信息配置的传输资源中的索引的指示信息。

5

10. 根据权利要求 1 至 9 任一项所述的方法，其特征在于，在所述接收第一信息之前，所述方法还包括：

接收第四信息，所述第四信息用于指示承载请求信息的资源，所述请求信息用于请求所述 N 个配置信息；

10 在所述承载请求信息的资源上发送所述请求信息。

11. 根据权利要求 10 所述的方法，其特征在于，所述请求信息包括第一正交序列，所述第一正交序列对应于所述本地信息。

15 12. 根据权利要求 1 至 11 任一项所述的方法，其特征在于，在所述接收第一信息之前，所述方法还包括：

接收第五信息，所述第五信息用于指示所述 N 个模型参数分别对应的本地信息，和/或，所述第五信息用于指示所述第一信息是否包括所述 N 个配置信息。

20 13. 一种通信方法，其特征在于，包括：

确定第一信息，所述第一信息指示 N 个配置信息，所述 N 个配置信息用于配置 N 个模型参数的资源，N 为正整数；其中，所述 N 个模型参数中的一个模型参数用于确定与第一节点的本地信息对应的神经网络模型；

发送所述第一信息。

25

14. 根据权利要求 13 所述的方法，其特征在于，所述本地信息包括本地资源开销水平，本地 AI 能力信息、本地的参数规模中的至少一项。

30 15. 根据权利要求 13 或 14 所述的方法，其特征在于，所述模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

16. 根据权利要求 13 至 15 任一项所述的方法，其特征在于，N 为大于 1 的整数，所述 N 个模型参数的参数规模不同。

35 17. 根据权利要求 13 至 15 任一项所述的方法，其特征在于，N 的取值为 1。

18. 根据权利要求 13 至 17 任一项所述的方法，其特征在于，所述方法还包括：

发送第二信息，所述第二信息指示 M 个配置信息，所述 M 个配置信息用于配置 M 个模型参数的传输资源，所述 M 个模型参数的参数规模不同，M 为大于 1 的整数；

40 基于所述 M 个配置信息发送所述 M 个模型参数的部分或全部。

19. 根据权利要求 18 所述的方法，其特征在于，所述 M 个配置信息中的任意一个配置信息包括以下至少一项：

指示所述一个配置信息配置的模型参数的参数规模的指示信息；

45 指示所述一个配置信息配置的传输资源的接收功率的指示信息；

指示所述一个配置信息配置的模型参数的参数类型的指示信息，所述参数类型包括梯度和/或权值；

指示所述一个配置信息配置的传输资源的频域资源的指示信息；

指示所述一个配置信息配置的传输资源的时域资源的指示信息。

20. 根据权利要求 13 至 19 任一项所述的方法，其特征在于，所述方法还包括：

5 发送第三信息，所述第三信息指示 K 个配置信息，所述 K 个配置信息用于配置 K 个模型参数的传输资源，所述 K 个模型参数的参数规模不同，K 为大于 1 的整数；

基于所述 K 个配置信息接收所述 K 个模型参数的部分或全部。

10 21. 根据权利要求 20 所述的方法，其特征在于，所述 K 个配置信息中的任意一个配置信息包括以下至少一项：

指示所述一个配置信息配置的模型参数的参数规模的指示信息；

指示所述一个配置信息配置的传输资源的发射功率的指示信息；

指示所述一个配置信息配置的模型参数的参数类型的指示信息，所述参数类型包括梯度和/或权值；

指示所述一个配置信息配置的传输资源的频域资源的指示信息；

15 指示所述一个配置信息配置的传输资源的时域资源的指示信息；

指示所述一个配置信息的传输资源在所述 K 个配置信息配置的传输资源中的索引的指示信息。

22. 根据权利要求 13 至 21 任一项所述的方法，其特征在于，在所述发送第一信息之前，所述方法还包括：

20 发送第四信息，所述第四信息用于指示承载请求信息的资源，所述请求信息用于请求所述 N 个配置信息；

在所述承载请求信息的资源上接收所述请求信息。

23. 根据权利要求 22 所述的方法，其特征在于，所述请求信息包括第一正交序列，所述第一正交序列对应于本地的参数规模。

24. 根据权利要求 13 至 23 任一项所述的方法，其特征在于，在所述接收第一信息之前，所述方法还包括：

30 发送第五信息，所述第五信息用于指示所述 N 个模型参数分别对应的本地信息，和/或，所述第五信息用于指示所述第一信息是否包括所述 N 个配置信息。

25. 一种通信装置，其特征在于，包括处理单元和收发单元；

所述收发单元用于接收第一信息，所述第一信息指示 N 个配置信息，所述 N 个配置信息用于配置 N 个模型参数的资源，N 为正整数；

35 所述处理单元用于确定所述 N 个配置信息中的第一配置信息；

所述收发单元还用于基于所述第一配置信息接收第一模型参数，所述第一模型参数用于确定与本地信息对应的第一神经网络模型。

40 26. 根据权利要求 25 所述的装置，其特征在于，所述本地信息包括本地资源开销水平，本地 AI 能力信息、本地的参数规模中的至少一项。

27. 根据权利要求 25 或 26 所述的装置，其特征在于，所述模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

45 28. 根据权利要求 25 至 27 任一项所述的装置，其特征在于，N 为大于 1 的整数，所述 N 个模型参数的参数规模不同，所述处理单元还用于基于所述本地信息在所述 N 个配置信息中确定所述第一配置信息。

29. 根据权利要求 25 至 27 任一项所述的装置，其特征在于，N 的取值为 1，其中，所述第一模型参数用于配置至少两个第一节点共用的神经网络模型；所述处理单元还用于基于所述本地信息将所述至少两个第一节点共用的神经网络模型进行知识蒸馏处理，得到所述第一神经网络模型。

5

30. 根据权利要求 25 至 29 任一项所述的装置，其特征在于，

所述收发单元还用于接收第二信息，所述第二信息指示 M 个配置信息，所述 M 个配置信息用于配置 M 个模型参数的传输资源，所述 M 个模型参数的参数规模不同，M 为大于 1 的整数；

所述处理单元还用于确定所述 M 个配置信息中的第二配置信息；

10 所述收发单元还用于基于所述第二配置信息接收第二模型参数，其中，所述第二模型参数的参数规模与本地的参数规模相同。

31. 根据权利要求 30 所述的装置，其特征在于，所述 M 个配置信息中的任意一个配置信息包括以下至少一项：

15 指示所述一个配置信息配置的模型参数的参数规模的指示信息；
指示所述一个配置信息配置的传输资源的接收功率的指示信息；
指示所述一个配置信息配置的模型参数的参数类型的指示信息，所述参数类型包括梯度和/或权值；
指示所述一个配置信息配置的传输资源的频域资源的指示信息；
指示所述一个配置信息配置的传输资源的时域资源的指示信息。

20

32. 根据权利要求 25 至 31 任一项所述的装置，其特征在于，

所述收发单元还用于接收第三信息，所述第三信息指示 K 个配置信息，所述 K 个配置信息用于配置 K 个模型参数的传输资源，所述 K 个模型参数的参数规模不同，K 为大于 1 的整数；

所述处理单元还用于确定所述 K 个配置信息中的第三配置信息；

25 所述收发单元还用于基于所述第三配置信息发送第三模型参数，其中，所述第三模型参数的参数规模与本地的参数规模相同。

33. 根据权利要求 32 所述的装置，其特征在于，所述 K 个配置信息中的任意一个配置信息包括以下至少一项：

30 指示所述一个配置信息配置的模型参数的参数规模的指示信息；
指示所述一个配置信息配置的传输资源的发射功率的指示信息；
指示所述一个配置信息配置的模型参数的参数类型的指示信息，所述参数类型包括梯度和/或权值；
指示所述一个配置信息配置的传输资源的频域资源的指示信息；
指示所述一个配置信息配置的传输资源的时域资源的指示信息；
35 指示所述一个配置信息的传输资源在所述 K 个配置信息配置的传输资源中的索引的指示信息。

34. 根据权利要求 25 至 33 任一项所述的装置，其特征在于，

所述收发单元还用于接收第四信息，所述第四信息用于指示承载请求信息的资源，所述请求信息用于请求所述 N 个配置信息；

40 所述收发单元还用于在所述承载请求信息的资源上发送所述请求信息。

35. 根据权利要求 34 所述的装置，其特征在于，所述请求信息包括第一正交序列，所述第一正交序列对应于本地的参数规模。

45

36. 根据权利要求 25 至 35 任一项所述的装置，其特征在于，

所述收发单元还用于接收第五信息，所述第五信息用于指示所述 N 个模型参数分别对应的本地信息，

和/或, 所述第五信息用于指示所述第一信息是否包括所述 N 个配置信息。

37. 一种通信装置, 其特征在于, 包括处理单元和收发单元;

所述处理单元用于确定第一信息, 所述第一信息指示 N 个配置信息, 所述 N 个配置信息用于配置 N 个模型参数的资源, N 为正整数; 其中, 所述 N 个模型参数中的一个模型参数用于确定与第一节点的本地信息对应的神经网络模型;

所述收发单元用于发送所述第一信息。

38. 根据权利要求 37 所述的装置, 其特征在于, 所述本地信息包括本地资源开销水平, 本地 AI 能力信息、本地的参数规模中的至少一项。

39. 根据权利要求 37 或 38 所述的装置, 其特征在于, 所述模型参数包括模型的权值、模型的激活函数、模型的偏置量中的至少一项。

40. 根据权利要求 37 至 39 任一项所述的装置, 其特征在于, N 为大于 1 的整数, 所述 N 个模型参数的参数规模不同。

41. 根据权利要求 37 至 39 任一项所述的装置, 其特征在于, N 的取值为 1。

42. 根据权利要求 37 至 41 任一项所述的装置, 其特征在于,

所述收发单元还用于发送第二信息, 所述第二信息指示 M 个配置信息, 所述 M 个配置信息用于配置 M 个模型参数的传输资源, 所述 M 个模型参数的参数规模不同, M 为大于 1 的整数;

所述收发单元还用于基于所述 M 个配置信息发送所述 M 个模型参数的部分或全部。

43. 根据权利要求 42 所述的装置, 其特征在于, 所述 M 个配置信息中的任意一个配置信息包括以下至少一项:

指示所述一个配置信息配置的模型参数的参数规模的指示信息;

指示所述一个配置信息配置的传输资源的接收功率的指示信息;

指示所述一个配置信息配置的模型参数的参数类型的指示信息, 所述参数类型包括梯度和/或权值;

指示所述一个配置信息配置的传输资源的频域资源的指示信息;

指示所述一个配置信息配置的传输资源的时域资源的指示信息。

44. 根据权利要求 37 至 43 任一项所述的装置, 其特征在于,

所述收发单元还用于发送第三信息, 所述第三信息指示 K 个配置信息, 所述 K 个配置信息用于配置 K 个模型参数的传输资源, 所述 K 个模型参数的参数规模不同, K 为大于 1 的整数;

所述收发单元还用于基于所述 K 个配置信息接收所述 K 个模型参数的部分或全部。

45. 根据权利要求 44 所述的装置, 其特征在于, 所述 K 个配置信息中的任意一个配置信息包括以下至少一项:

指示所述一个配置信息配置的模型参数的参数规模的指示信息;

指示所述一个配置信息配置的传输资源的发射功率的指示信息;

指示所述一个配置信息配置的模型参数的参数类型的指示信息, 所述参数类型包括梯度和/或权值;

指示所述一个配置信息配置的传输资源的频域资源的指示信息;

指示所述一个配置信息配置的传输资源的时域资源的指示信息;

指示所述一个配置信息的传输资源在所述 K 个配置信息配置的传输资源中的索引的指示信息。

5

46. 根据权利要求 37 至 45 任一项所述的装置，其特征在于，
所述收发单元还用于发送第四信息，所述第四信息用于指示承载请求信息的资源，所述请求信息用于请求所述 N 个配置信息；
所述收发单元还用于在所述承载请求信息的资源上接收所述请求信息。

10

47. 根据权利要求 46 所述的装置，其特征在于，所述请求信息包括第一正交序列，所述第一正交序列对应于本地的参数规模。

48. 根据权利要求 37 至 47 任一项所述的装置，其特征在于，
所述收发单元还用于发送第五信息，所述第五信息用于指示所述 N 个模型参数分别对应的本地信息，和/或，所述第五信息用于指示所述第一信息是否包括所述 N 个配置信息。

15

49. 一种通信装置，其特征在于，包括至少一个处理器，用于执行存储器中的计算机程序或指令，以实现如权利要求 1 至 12 中任一项所述的方法。

50. 一种通信装置，其特征在于，包括至少一个处理器，用于执行存储器中的计算机程序或指令，以实现如权利要求 13 至 24 中任一项所述的方法。

20

51. 一种通信系统，其特征在于，所述系统包括如权利要求 25 至 36 任一项所述的通信装置，以及，用于执行如权利要求 37 至 48 任一项所述的通信装置；或，
所述系统包括如权利要求 49 所述的通信装置，以及，如权利要求 50 所述的通信装置。

25

52. 一种计算机可读存储介质，其特征在于，所述介质存储有指令，当所述指令被计算机执行时，实现权利要求 1 至 24 中任一项所述的方法。

53. 一种计算机程序产品，其特征在于，包括指令，当所述指令在计算机上运行时，使得计算机执行如权利要求 1 至 24 中任一项所述的方法。

30