

一种信道接入方法及相关装置

本申请要求于2021年6月17日提交中国国家知识产权局、申请号为202110673131.6、申请名称为“一种信道接入方法及相关装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及通信技术领域，尤其涉及一种信道接入方法及相关装置。

背景技术

在短距离/无线局域网(Wireless Fidelity, Wi-Fi)等无线网络中，传输数据的信道是共享的。那么，如果在特定区域内存在多个站点(station, STA)同时向同一接入点(access point, AP)发送报文，就会发生冲突，进而导致数据传输失败。

目前，Wi-Fi系统采用载波侦听多路访问/冲突避免(carrier sense multiple access / collision avoidance, CSMA/CA)机制避免在共享信道上产生冲突。即具有侦听能力的STA在报文到达时，侦听随机时长内的信道状态，若该随机时长内信道处于空闲状态，则该STA接入信道。

上述采用CSMA/CA机制避免共享信道的冲突方式，可被认为是一种冲突分解类算法，即期望通过完全随机化达到冲突分解的效果。也就是说，该方式中的每个STA不具有预测其他STA是否接入信道的能力，从而导致系统的吞吐量较低、时延较高。

发明内容

本申请实施例提供了一种信道接入方法及相关装置，有利于提高系统的吞吐量和降低时延。

第一方面，本申请实施例提供一种信道接入方法。该方法中，接入点AP接收N个站点STA分别上报的N个动作信息，N个动作信息用于确定每个STA的第一神经网络的训练结果，进而AP根据该N个动作信息，确定每个STA的第一神经网络的训练结果，并将每个STA的第一神经网络的训练结果发送给对应的STA。

可见，每个STA的第一神经网络的训练结果是根据N个STA上报的动作信息确定的，而不是只根据自身STA的动作信息确定的，可提高第一神经网络的预测能力，从而有利于提高STA对是否信道接入的预测能力，进而有利于提高系统的吞吐量和降低时延。

一种可选的实施方式中，动作信息是指一段时间的动作，该动作为发送或不发送。该一段时间是指STA上一次成功上报动作信息的时刻距离当前时刻的时间。也就是说，动作是指STA自上一次成功上报动作信息以来，发送或不发送报文的动作。

一种可选的实施方式中，AP还可接收N个STA分别上报的载波侦听结果信息或报文传输结果信息。载波侦听结果信息包括载波侦听结果，报文传输结果信息包括报文传输结果。从而，AP根据N个动作信息，确定每个STA的第一神经网络的训练结果，是指AP根据N个动作信息和N个载波侦听结果信息，确定每个STA的第一神经网络的训练结果；或者，是指AP根据N个动作信息和N个报文传输结果信息，确定每个STA的第一神经网络的训练结果。

果。

可见，每个 STA 还可向 AP 上报载波侦听结果信息或报文传输结果信息。从而 AP 可直接根据 N 个动作信息和 N 个载波侦听结果信息，训练每个 STA 的第一神经网络，或者根据 N 个动作信息和 N 个报文传输结果信息，训练每个 STA 的第一神经网络，进而有利于降低 AP 处理的复杂度。

一种可选的实施方式中，训练结果为神经网络参数或梯度，神经网络参数/梯度用于对应的 STA 更新第一神经网络。

一种可选的实施方式中，当 AP 接收的是 N 个 STA 分别上报的动作信息时，动作信息携带于 STA 上报的第一帧的动作详细信息字段中，动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段，T 为正整数。

其中，时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息，即第一响应信息是该 STA 上次一成功上报动作信息时接收的响应信息，响应信息可以是确认信息 ACK。数据 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个时隙内的动作，即数据 1 子字段用于指示 STA 在上一次成功上报动作信息后的第一个时隙上动作。数据 T 子字段用于指示 STA 上一次成功接收第一响应信息后的第 T 个时隙内的动作，该第 T 个时隙也为 STA 当前上报动作信息前的最后一个时隙。

可见，针对 N 个 STA，每个 STA 上报的动作信息携带于第一帧中，且每个 STA 向 AP 上报的动作信息包括 STA 上一次成功上报动作信息的时刻，以及在上一次成功上报动作信息后的第一个时隙至第 T 个时隙内的动作。

另一种可选的实施方式中，当 AP 接收的是 N 个 STA 分别上报的动作信息时，动作信息携带于 STA 上报的第一帧的动作详细信息字段中。动作详细信息字段包括时间指示子字段、动作 1 子字段和时间 1 子字段、……、动作 P 子字段和时间 P 子字段，P 为正整数。

其中，时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。也就是说，时间指示子字段是用于指示 STA 上一次成功上报动作信息的时刻。

动作 1 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的第一个动作。动作 P 子字段用于指示 STA 在上一次成功接收所述第一响应信息的时刻距离当前时刻内的第 P 个动作。也就是说，动作 1 子字段是用于指示 STA 在上一次成功上报动作信息后的第一个动作，动作 P 子字段是用于指示 STA 在上一次成功上报动作信息后，当前时刻前的最后一个动作。

时间 1 子字段用于指示动作 1 的持续时间或动作 1 的结束时间。时间 P 子字段用于指示动作 P 的持续时间或动作 P 的结束时间。当时间 1 子字段用于指示动作 1 的持续时间，时间 P 子字段用于指示动作 P 的持续时间时，动作不同，其持续时间代表的含义不一样。当动作为发送动作时，持续时间代表的是发送的报文的包长。当动作为不发送动作时，持续时间代表的是不发送报文的持续时间。

可见，针对 N 个 STA，每个 STA 上报的动作信息携带于第一帧中，且每个 STA 向 AP 上报的动作信息包括 STA 上一次成功上报动作信息的时刻，STA 在上一次成功上报动作信息后的每个动作，以及每个动作的持续时间或结束时间。

又一种可选的实施方式中，当 AP 接收的是 N 个 STA 分别上报的动作信息时，动作信息携带于 STA 上报的第一帧的动作详细信息字段中。动作详细信息字段包括时间 1 指示子字段和动作 1 子字段、……、时间 P 指示子字段和动作 P 子字段，P 为正整数。

其中，动作 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个动作。动作 P 子字段用于指示 STA 在上一次成功接收第一响应信息后的时刻距离当前时刻内的第 P 个动作。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。也就是说，动作 1 子字段是用于指示 STA 在上一次成功上报动作信息后的第一个动作，动作 P 子字段是用于指示 STA 在上一次成功上报动作信息后，当前时刻前的最后一个动作。时间 1 指示子字段用于指示动作 1 的起始时间。时间 P 指示子字段用于指示动作 P 的起始时间。

可见，针对 N 个 STA，每个 STA 上报的动作信息携带于第一帧中，且每个 STA 向 AP 上报的动作信息包括 STA 上一次成功上报动作信息后的每个动作，以及每个动作的起始时间。

又一种可选的实施方式中，当 AP 接收的是 N 个 STA 分别上报的动作信息时，动作信息携带于 STA 上报的第一帧的动作详细信息字段中。动作详细信息字段包括时间 1 指示子字段和持续时间 1 子字段、……、时间 K 指示子字段和持续时间 K 子字段，K 为正整数。

其中，时间 1 指示子字段用于指示动作 1 的开始时间/结束时间。动作 1 是 STA 在上一次成功接收第一响应信息后，第一次发送报文且未接收到第二响应信息时的发送动作。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。第二响应信息是 AP 成功接收 STA 发送的报文时发送的响应信息。持续时间 1 子字段用于指示动作 1 的持续时间。

时间 K 指示子字段用于指示动作 K 的开始时间/结束时间。动作 K 是 STA 在上一次成功接收第一响应信息后，第 K 次发送报文且未接收到第二响应信息时的发送动作。持续时间 K 子字段用于指示动作 K 的持续时间。

可见，针对 N 个 STA，每个 STA 上报的动作信息携带于第一帧中，且每个 STA 向 AP 上报的动作信息包括 STA 上一次成功上报动作信息后的每次发送报文且未成功时，发送动作的开始时间/结束时间，以及每次发送报文不成功时发生的报文的时长。

又一种可选的实施方式中，当 AP 接收的是 N 个 STA 分别上报的动作信息时，动作信息携带于 STA 上报的第一帧的动作详细信息字段中。动作详细信息字段包括第一时间 1 指示子字段和第二时间 1 指示子字段、……、第一时间 K 指示子字段和第二时间 K 指示子字段，K 为正整数。

其中，第一时间 1 指示子字段用于指示动作 1 的起始时间。第一时间 K 指示子字段用于指示动作 K 的起始时间。动作 1 是 STA 在上一次成功接收第一响应信息后，第一次发送报文且未接收到第二响应信息时的发送动作。动作 K 是 STA 在上一次成功接收所述第一响应信息后，第 K 次发送报文且未接收到所述第二响应信息时的发送动作。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。第二响应信息是 AP 成功接收 STA 发送的报文时发送的响应信息。也就是说，动作 1 是对应的 STA 在上一次成功上报动作信息后，第一次发送报文不成功的动作，动作 K 是该 STA 在上一次成功上报动作信息后，第 K 次发送报文不成功的动作。

第二时间 1 指示子字段用于指示动作 1 的结束时间。第二时间 K 指示子字段用于指示动作 K 的结束时间。

可见，针对 N 个 STA，每个 STA 上报的动作信息携带于第一帧中，且每个 STA 向 AP 上报的动作信息包括 STA 上一次成功上报动作信息后每次发送报文且未成功时发送动作的开始时间和结束时间。

又一种可选的实施方式中，当 AP 接收到的是 N 个 STA 分别上报的动作信息和载波侦听结果信息时，动作信息和载波侦听结果信息携带于 STA 上报的第一帧的动作详细信息字段中。动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段，T 为正整数。

其中，时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。

数据 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个时隙内的动作和载波侦听结果。数据 T 子字段用于指示 STA 上一次成功接收第一响应信息后的第 T 个时隙内的动作和载波侦听结果。

可见，针对 N 个 STA，每个 STA 上报的动作信息和载波侦听结果信息携带于第一帧中，且每个 STA 向 AP 上报的信息包括 STA 上一次成功上报动作信息的时刻，STA 自上一次成功上报动作信息以来每个时隙的动作和载波侦听结果。

又一种可选的实施方式中，当 AP 接收到的是 N 个 STA 分别上报的动作信息和报文传输结果信息时，动作信息和报文传输结果信息携带于 STA 上报的第一帧的动作详细信息字段中。动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段，T 为正整数。

其中，时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。

数据 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个时隙内的动作和报文传输结果。数据 T 子字段用于指示 STA 上一次成功接收第一响应信息后的第 T 个时隙内的动作和报文传输结果。

可见，针对 N 个 STA，每个 STA 上报的动作信息和报文传输结果信息携带于第一帧中，且每个 STA 向 AP 上报的信息包括 STA 上一次成功上报动作信息的时刻，STA 自上一次成功上报动作信息以来每个时隙的动作和报文传输结果。

一种可选的实施方式中，AP 根据 N 个动作信息，确定每个 STA 的第一神经网络的训练结果，是指：AP 将每个 STA 的状态信息输入到对应 STA 的第一神经网络，获得第一神经网络的输出值；AP 将每个第一神经网络的输出值输入到第二神经网络，获得第二神经网络的输出值，第二神经网络的输出值用于表征预设时间内的期望奖励；AP 根据第二神经网络的输出值和奖励函数训练第三神经网络，通过最小化所述第三神经网络的损失函数，确定每个第一神经网络的训练结果，第三神经网络包括每个第一神经网络和第二神经网络。

其中，STA 的状态信息是根据 STA 的动作信息获得的，第二神经网络的神经网络参数是根据 N 个动作信息获得的，奖励函数是根据 N 个动作信息确定的；或者，

STA 的状态信息是根据 STA 的动作信息和载波侦听结果信息获得的，第二神经网络的神经网络参数是根据 N 个动作信息和 N 个载波侦听结果信息获得的，奖励函数是根据 N 个动作信息和 N 个载波侦听结果信息确定的；或者，

STA 的状态信息是根据 STA 的动作信息和报文传输结果信息获得的，第二神经网络的神经网络参数是根据 N 个动作信息和 N 个报文传输结果信息获得的，奖励函数是根据 N 个动作信息和 N 个报文传输结果信息确定的。

可见，AP 是先将根据每个 STA 上报的信息获得的状态信息输入到该 STA 的第一神经网络中，获得每个第一神经网络的输出值，再将 N 个第一神经网络的输出值输入到第二神经网络中，获得第二神经网络的输出值，然后根据损失函数训练第三神经网络，最终获得第一神经网络的训练结果。每个 STA 的第一神经网络的训练结果是根据 N 个 STA 上报的信息确定的，而不是只根据自身 STA 的信息确定的，从而有利于提高每个 STA 对其他 STA 信道接入行为的预测能力。

一种可选的实施方式中，AP 还可在根据 N 个动作信息确定第一 STA 发送报文成功时，将奖励函数的值设置为 1。第一 STA 是 N 个 STA 中上一次成功接收第二响应信息的时刻距

离当前时刻的时间间隔最长的 STA。

可见, AP 在根据 N 个 STA 上报的信息确定距离上一次成功发送报文以来时间间隔最长的 STA 发送报文成功时, 将奖励函数的值设置为 1。

又一种可选的实施方式中, AP 还可在根据 N 个动作信息确定第二 STA 发送报文成功时, 将奖励函数的值设置为第一时长减 1。第二 STA 是 N 个 STA 中, 除第一 STA 外的 STA, 第一 STA 是 N 个 STA 中上一次成功接收第二响应信息的时刻距离当前时刻的时间间隔最长的 STA。第一时长是第二 STA 上一次成功接收第二响应信息的时刻距离当前时刻的时长。

可见, AP 在根据 N 个 STA 上报的信息确定除距离上一次成功发送报文以来时间间隔最长的 STA 之外的 STA 发送报文成功时, 将奖励函数的值设置为该 STA 距离上一次成功发送报文的时间间隔减去 1。

一种可选的实施方式中, AP 还可在根据 N 个动作信息确定 N 个 STA 中的 M 个 STA 在同一时隙上发送报文时, 将奖励函数的值设置为 -1。M 为小于或等于 N 的正整数。可见, AP 在根据 N 个 STA 上报的信息确定 N 个 STA 中的部分 STA 在同一时隙上发送报文时, 将奖励函数设置为 -1。

一种可选的实施方式中, AP 还可在根据 N 个动作信息确定 N 个 STA 在同一时隙上均未发送报文时, 将奖励函数的值设置为 0。可见, AP 在根据 N 个 STA 上报的信息确定 N 个 STA 在同一时隙上均未发送报文时, 将奖励函数的值设置为 0。

一种可选的实施方式中, N 个 STA 共享神经网络参数。那么, AP 将每个 STA 的第一神经网络的训练结果发送给对应的 STA, 是指: AP 将第一神经网络的训练结果广播给 N 个 STA。可见, 当 N 个 STA 共享神经网络参数时, AP 根据 N 个 STA 上报的信息, 训练每个第一神经网络获得的训练结果相同, 则 AP 可通过广播方式将训练结果告知给每个 STA, 从而可减少系统的信令开销。

一种可选的实施方式中, N 个 STA 中的 S 个 STA 共享神经网络参数, S 为小于或等于 N 的正整数, AP 将每个 STA 的第一神经网络的训练结果发送给对应的 STA, 是指: AP 将 S 个 STA 对应的第一神经网络的训练结果组播给 S 个 STA, 以及将 N-S 个第一神经网络的训练结果单播给对应的 STA。可见, 当 N 个 STA 中的部分 STA 共享神经网络参数时, AP 可将共享的神经网络参数对应的训练结果通过组播方式告知该部分的 STA, 通过单播的方式将未共享神经网络参数对应的训练结果单播给这些 STA。该方式中, 通过组播告知共享一个神经网络的 STA 的训练结果, 也可减少系统的开销。

一种可选的实施方式中, N 个 STA 不共享神经网络参数时, 每个第一神经网络的训练结果是单播给对应的 STA 的。

第二方面, 本申请还提供了一种信道接入方法。该方面的信道接入方法与第一方面所述的信道接入方法相对应, 该方面的信道接入方法是从站点 STA 侧进行阐述的。该方法中, 站点 STA 向接入点 AP 上报动作信息, 动作信息用于确定第一神经网络的训练结果, 第一神经网络为 STA 的神经网络; STA 接收来自 AP 的第一神经网络的训练结果, 第一神经网络的训练结果是基于所述动作信息获得的, 第一神经网络的训练结果用于更新第一神经网络, 以确定 STA 是否接入信道; STA 根据第一神经网络的训练结果更新第一神经网络, 并在侦听到信道空闲时根据更新后的第一神经网络和当前的状态信息确定是否接入信道。

可见, 本申请实施例中, STA 向 AP 上报动作信息, 并接收 AP 根据该动作信息训练第一神经网络获得的训练结果, 从而 STA 根据该训练结果更新第一神经网络, 并在侦听信道空闲时, 根据更新后的第一神经网络和侦听到的动作信息确定是否接入信道。用于更新每个第

一神经网络的训练结果是 AP 根据 N 个 STA 上报的动作信息确定的,从而第一神经网络的预测性更好,进而 STA 根据更新后的第一神经网络自行确定是否接入信道时,确定接入信道或不接入信道的准确率更好,可提高通信系统的吞吐量,降低通信时延。

一种可选的实施方式中,STA 还可向 AP 上报载波侦听结果信息或报文传输结果信息,载波侦听结果信息或报文传输结果信息用于确定第一神经网络的训练结果。可见,STA 除了向 AP 上报动作信息外,还可向 AP 上报载波侦听结果信息或报文传输结果信息,从而有利于 AP 直接根据 N 个 STA 上报的信息训练第一神经网络,进而有利于降低 AP 处理的复杂度。

一种可选的实施方式中,训练结果为神经网络参数或梯度;载波侦听结果信息或报文传输结果信息用于确定第一神经网络的训练结果。

一种可选的实施方式中,当 STA 上报的是动作信息时,动作信息携带于第一帧的动作详细信息字段中。动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段,T 为正整数。

其中,时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息,即第一响应信息是该 STA 上次一成功上报动作信息时接收的响应信息,响应信息可以是确认信息 ACK。数据 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个时隙内的动作,即数据 1 子字段用于指示 STA 在上一次成功上报动作信息后的第一个时隙上动作。数据 T 子字段用于指示 STA 上一次成功接收第一响应信息后的第 T 个时隙内的动作,该第 T 个时隙也为 STA 当前上报动作信息前的最后一个时隙。

可见,STA 上报的动作信息携带于第一帧中,且 STA 向 AP 上报的动作信息包括 STA 上一次成功上报动作信息的时刻,以及在上一次成功上报动作信息后的第一个时隙至第 T 个时隙内的动作。

另一种可选的实施方式中,当 STA 上报的是动作信息时,动作信息携带于 STA 上报的第一帧的动作详细信息字段中。作详细信息字段包括时间指示子字段、动作 1 子字段和时间 1 子字段、……、动作 P 子字段和时间 P 子字段,P 为正整数。

其中,时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。也就是说,时间指示子字段是用于指示 STA 上一次成功上报动作信息的时刻。

动作 1 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的第一个动作。动作 P 子字段用于指示 STA 在上一次成功接收所述第一响应信息的时刻距离当前时刻内的第 P 个动作。也就是说,动作 1 子字段是用于指示 STA 在上一次成功上报动作信息后的第一个动作,动作 P 子字段是用于指示 STA 在上一次成功上报动作信息后,距离当前时刻前的最后一个动作。

时间 1 子字段用于指示动作 1 的持续时间或动作 1 的结束时间。时间 P 子字段用于指示动作 P 的持续时间或动作 P 的结束时间。当时间 1 子字段用于指示动作 1 的持续时间,时间 P 子字段用于指示动作 P 的持续时间时,动作不同,其持续时间代表的含义不一样。当动作为发送动作时,持续时间代表的是发送的报文的包长。当动作为不发送动作时,持续时间代表的是不发送报文的持续时间。

可见,STA 上报的动作信息携带于第一帧中,且 STA 向 AP 上报的动作信息包括 STA 上一次成功上报动作信息的时刻,STA 在上一次成功上报动作信息后的每个动作,以及每个动作的持续时间或结束时间。

又一种可选的实施方式中，当 STA 上报的是动作信息时，动作信息携带于 STA 上报的第一帧的动作详细信息字段中。动作详细信息字段包括时间 1 指示子字段和动作 1 子字段、……、时间 P 指示子字段和动作 P 子字段，P 为正整数。

动作 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个动作。动作 P 子字段用于指示 STA 在上一次成功接收第一响应信息后的时刻距离当前时刻内的第 P 个动作。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。也就是说，动作 1 子字段是用于指示 STA 在上一次成功上报动作信息后的第一个动作，动作 P 子字段是用于指示 STA 在上一次成功上报动作信息后，当前时刻前的最后一个动作。时间 1 指示子字段用于指示动作 1 的起始时间。时间 P 指示子字段用于指示动作 P 的起始时间。

可见，STA 上报的动作信息携带于第一帧中，且 STA 向 AP 上报的动作信息包括 STA 上一次成功上报动作信息后的每个动作，以及每个动作的起始时间。

又一种可选的实施方式中，当 STA 上报的是动作信息时，动作信息携带于 STA 上报的第一帧的动作详细信息字段中。动作详细信息字段包括时间 1 指示子字段和持续时间 1 子字段、……、时间 K 指示子字段和持续时间 K 子字段，K 为正整数。

其中，时间 1 指示子字段用于指示动作 1 的开始时间/结束时间。动作 1 是 STA 在上一次成功接收第一响应信息后，第一次发送报文且未接收到第二响应信息时的发送动作。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。第二响应信息是 AP 成功接收 STA 发送的报文时发送的响应信息。持续时间 1 子字段用于指示动作 1 的持续时间。

时间 K 指示子字段用于指示动作 K 的开始时间/结束时间。动作 K 是 STA 在上一次成功接收第一响应信息后，第 K 次发送报文且未接收到第二响应信息时的发送动作。持续时间 K 子字段用于指示动作 K 的持续时间。

可见，STA 上报的动作信息携带于第一帧中，且 STA 向 AP 上报的动作信息包括 STA 上一次成功上报动作信息后的每次发送报文且未成功时，发送动作的开始时间/结束时间，以及每次发送报文不成功时报文的时长。

又一种可选的实施方式中，当 STA 上报的是动作信息时，动作信息携带于 STA 上报的第一帧的动作详细信息字段中。动作详细信息字段包括第一时间 1 指示子字段和第二时间 1 指示子字段、……、第一时间 K 指示子字段和第二时间 K 指示子字段，K 为正整数。

其中，第一时间 1 指示子字段用于指示动作 1 的起始时间。第一时间 K 指示子字段用于指示动作 K 的起始时间。动作 1 是 STA 在上一次成功接收第一响应信息后，第一次发送报文且未接收到第二响应信息时的发送动作。动作 K 是 STA 在上一次成功接收所述第一响应信息后，第 K 次发送报文且未接收到所述第二响应信息时的发送动作。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。第二响应信息是 AP 成功接收 STA 发送的报文时发送的响应信息。也就是说，动作 1 是对应的 STA 在上一次成功上报动作信息后，第一次发送报文不成功的动作，动作 K 是该 STA 在上一次成功上报动作信息后，第 K 次发送报文不成功的动作。

第二时间 1 指示子字段用于指示动作 1 的结束时间。第二时间 K 指示子字段用于指示动作 K 的结束时间。

可见，STA 上报的动作信息携带于第一帧中，且 STA 向 AP 上报的动作信息包括 STA 上一次成功上报动作信息后每次发送报文且未成功时发送动作的开始时间和结束时间。

又一种可选的实施方式中，当 STA 上报的是动作信息和载波侦听结果信息时，动作信息和载波侦听结果信息携带于 STA 上报的第一帧的动作详细信息字段中。动作详细信息字段包

括时间指示子字段、数据 1 子字段至数据 T 子字段, T 为正整数。

其中, 时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。

数据 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个时隙内的动作和载波侦听结果。数据 T 子字段用于指示 STA 上一次成功接收第一响应信息后的第 T 个时隙内的动作和载波侦听结果。

可见, STA 上报的动作信息和载波侦听结果信息携带于第一帧中, 且 STA 向 AP 上报的信息包括 STA 上一次成功上报动作信息的时刻, STA 自上一次成功上报动作信息以来每个时隙的动作和载波侦听结果。

又一种可选的实施方式中, 当 STA 上报的是动作信息和报文传输结果信息时, 动作信息和报文传输结果信息携带于 STA 上报的第一帧的动作详细信息字段中。动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段, T 为正整数。

其中, 时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。

数据 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个时隙内的动作和报文传输结果。数据 T 子字段用于指示 STA 上一次成功接收第一响应信息后的第 T 个时隙内的动作和报文传输结果。

可见, STA 上报的动作信息和报文传输结果信息携带于第一帧中, 且 STA 向 AP 上报的信息包括 STA 上一次成功上报动作信息的时刻, STA 自上一次成功上报动作信息以来每个时隙的动作和报文传输结果。

一种可选的实施方式中, STA 根据第一神经网络的训练结果更新第一神经网络, 并在侦听到信道空闲时根据更新后的第一神经网络和该 STA 当前的状态信息确定是否接入信道, 是指: STA 将该 STA 当前的状态信息输入到更新后的第一神经网络, 输出第一值和第二值, 第一值用于表征接入信道获得的预期奖励, 第二值用于表征不接入信道获得的预期奖励; 然后, STA 在第一值大于第二值时, 确定接入信道; STA 在第一值小于第二值时, 确定不接入信道。

可见, STA 在侦听到信道空闲时, 将侦听到的动作信息输入到更新后的第一神经网络, 获得接入信道的期望奖励和不接入信道的期望奖励, 并在接入信道的期望奖励大于不接入信道的期望奖励时, 确定接入信道。

第三方面, 本申请还提供一种通信装置。该通信装置具有实现上述第一方面所述的 AP 的部分或全部功能, 或者具有实现上述第二方面所述的 STA 的部分或全部功能。比如, 该通信装置的功能可具备本申请中第一方面所述的 AP 的部分或全部实施例中的功能, 也可以具备单独实施本申请中的任一个实施例的功能。所述功能可以通过硬件实现, 也可以通过硬件执行相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的单元或模块。

在一种可能的设计中, 该通信装置的结构中可包括处理单元和通信单元, 所述处理单元被配置为支持通信装置执行上述方法中相应的功能。所述通信单元用于支持通信装置与其他通信装置之间的通信。所述通信装置还可以包括存储单元, 所述存储单元用于与处理单元和收发单元耦合, 其保存通信装置必要的程序指令和数据。

一种实施方式中, 所述通信装置包括:

通信单元, 用于接收 N 个站点 STA 上报的动作信息, N 个动作信息用于确定每个 STA 的第一神经网络的训练结果, N 为正整数;

处理单元, 用于根据 N 个动作信息, 确定每个 STA 的第一神经网络的训练结果;

通信单元，还用于将每个 STA 的第一神经网络的训练结果发送给对应的 STA。

另外，该方面中，通信装置其他可选的实施方式可参见上述第一方面的相关内容，此处不再详述。

另一种实施方式中，所述通信装置包括：

通信单元，用于向接入点 AP 上报动作信息，动作信息用于确定处理单元的第一神经网络的训练结果；

通信单元，还用于接收来自 AP 的第一神经网络的训练结果，第一神经网络的训练结果用于更新第一神经网络，以确定处理单元是否接入信道；

处理单元，用于根据第一神经网络的训练结果更新第一神经网络，并在侦听到信道空闲时根据更新后的第一神经网络和处理单元当前的状态信息确定是否接入信道。

另外，该方面中，通信装置其他可选的实施方式可参见上述第二方面的相关内容，此处不再详述。

作为示例，收发单元可以为收发器或通信接口，存储单元可以为存储器，处理单元可以为处理器。

一种实施方式中，所述通信装置包括：

收发器，用于接收 N 个站点 STA 上报的动作信息，N 个动作信息用于确定每个 STA 的第一神经网络的训练结果，N 为正整数；

处理器，用于根据 N 个动作信息，确定每个 STA 的第一神经网络的训练结果；

收发器，还用于将每个 STA 的第一神经网络的训练结果发送给对应的 STA。

另外，该方面中，上行通信装置其他可选的实施方式可参见上述第一方面的相关内容，此处不再详述。

另一种实施方式中，所述通信装置包括：

收发器，用于向接入点 AP 上报动作信息，动作信息用于确定处理器的第一神经网络的训练结果；

收发器，还用于接收来自 AP 的第一神经网络的训练结果，第一神经网络的训练结果用于更新第一神经网络，以确定处理器是否接入信道；

处理器，用于根据第一神经网络的训练结果更新第一神经网络，并在侦听到信道空闲时根据更新后的第一神经网络和处理器当前的状态信息确定是否接入信道。

另外，该方面中，通信装置其他可选的实施方式可参见上述第二方面的相关内容，此处不再详述。

另一种实施方式中，该通信装置为芯片或芯片系统。所述处理单元也可以体现为处理电路或逻辑电路；所述收发单元可以是该芯片或芯片系统上的输入/输出接口、接口电路、输出电路、输入电路、管脚或相关电路等。

在实现过程中，处理器可用于进行，例如但不限于，基带相关处理，收发器可用于进行，例如但不限于，射频收发。上述器件可以分别设置在彼此独立的芯片上，也可以至少部分的或者全部的设置在同一块芯片上。例如，处理器可以进一步划分为模拟基带处理器和数字基带处理器。其中，模拟基带处理器可以与收发器集成在同一块芯片上，数字基带处理器可以设置在独立的芯片上。随着集成电路技术的不断发展，可以在同一块芯片上集成的器件越来越多。例如，数字基带处理器可以与多种应用处理器（例如但不限于图形处理器，多媒体处理器等）集成在同一块芯片之上。这样的芯片可以称为系统芯片（System on a Chip, SoC）。将各个器件独立设置在不同的芯片上，还是整合设置在一个或者多个芯片上，往往取决于产

品设计的需要。本申请实施例对上述器件的实现形式不做限定。

第四方面，本申请还提供一种处理器，用于执行上述各种方法。在执行这些方法的过程中，上述方法中有关发送上述信息和接收上述信息的过程，可以理解为由处理器输出上述信息的过程，以及处理器接收输入的上述信息的过程。在输出上述信息时，处理器将该上述信息输出给收发器，以便由收发器进行发射。该上述信息在由处理器输出之后，还可能需要进行其他的处理，然后才到达收发器。类似的，处理器接收输入的上述信息时，收发器接收该上述信息，并将其输入处理器。更进一步的，在收发器收到该上述信息之后，该上述信息可能需要进行其他的处理，然后才输入处理器。

基于上述原理，举例来说，前述方法中提及的上报动作信息可以理解为处理器输出动作信息。

对于处理器所涉及的发射、发送和接收等操作，如果没有特殊说明，或者，如果未与其在相关描述中的实际作用或者内在逻辑相抵触，则均可以更加一般性的理解为处理器输出和接收、输入等操作，而不是直接由射频电路和天线所进行的发射、发送和接收操作。

在实现过程中，上述处理器可以是专门用于执行这些方法的处理器，也可以是执行存储器中的计算机指令来执行这些方法的处理器，例如通用处理器。上述存储器可以为非瞬态性（non-transitory）存储器，例如只读存储器（Read Only Memory, ROM），其可以与处理器集成在同一块芯片上，也可以分别设置在不同的芯片上，本申请实施例对存储器的类型以及存储器与处理器的设置方式不做限定。

第五方面，本申请还提供了一种通信系统，该系统包括上述方面的至少一个 AP 以及至少两个 STA。在另一种可能的设计中，该系统还可以包括本申请提供的方案中与 AP、STA 进行交互的其他设备。

第六方面，本申请提供了一种计算机可读存储介质，用于储存指令，当所述指令被通信装置执行时，实现上述第一方面、第二方面任一项所述的方法。

第七方面，本申请还提供了一种包括指令的计算机程序产品，当其在通信装置上运行时，使得通信装置执行上述第一方面、第二方面任一项所述的方法。

第八方面，本申请提供了一种芯片系统，该芯片系统包括处理器和接口，所述接口用于获取程序或指令，所述处理器用于调用所述程序或指令以实现或者支持 AP 实现第一方面所涉及的功能，或者用于调用所述程序或指令以实现或者支持 STA 实现第二方面所涉及的功能。例如，确定或处理上述方法中所涉及的数据和信息中的至少一种。在一种可能的设计中，所述芯片系统还包括存储器，所述存储器，用于保存终端必要的程序指令和数据。该芯片系统，可以由芯片构成，也可以包括芯片和其他分立器件。

第九方面，本申请提供一种通信装置，包括处理器，用于执行存储器中存储的计算机程序或可执行指令，当计算机程序或可执行指令被执行时，使得该装置执行如第一方面及第一方面各个可能的实现中的方法。

在一种可能的实现中，处理器和存储器集成在一起；

在另一种可能的实现中，上述存储器位于该通信装置之外。

第十方面，本申请提供一种通信装置，包括处理器，用于执行存储器中存储的计算机程序或可执行指令，当计算机程序或可执行指令被执行时，使得该装置执行如第二方面及第二方面各个可能的实现中的方法。

在一种可能的实现中，处理器和存储器集成在一起；

在另一种可能的实现中，上述存储器位于该通信装置之外。

附图说明

- 图 1 是本申请实施例提供的一种通信系统的结构示意图；
图 2 是本申请实施例提供的一种采用 CSMA/CA 机制避免信道冲突的示意图；
图 3 是本申请实施例提供的一种信道接入方法的交互示意图；
图 4 是本申请实施例提供的一种帧结构示意图；
图 5 是本申请实施例提供的另一种帧结构示意图；
图 6 (a) 是本申请实施例提供的一种训练数据单元格式示意图；
图 6 (b) 是本申请实施例提供的另一种训练数据单元格式示意图；
图 6 (c) 是本申请实施例提供的又一种训练数据单元格式示意图；
图 6 (d) 是本申请实施例提供的又一种训练数据单元格式示意图；
图 6 (e) 是本申请实施例提供的又一种训练数据单元格式示意图；
图 7 是本申请实施例提供的一种神经网络的结构示意图；
图 8 是本申请实施例提供的一种侦听信道的示意图；
图 9 是本申请实施例提供的一种神经网络训练的结构意图；
图 10 是本申请实施例提供的一种神经网络的结构意图；
图 11 是本申请实施例提供的一种信道接入方法的实现框图；
图 12 是本申请实施例提供的一种系统的吞吐量对比示意图；
图 13 是本申请实施例提供的一种系统的平均时延对比示意图；
图 14 是本申请实施例提供的一种系统的时延抖动对比示意图；
图 15 是本申请实施例提供的又一种帧结构示意图；
图 16 是本申请实施例提供的一种通信装置的结构示意图；
图 17 是本申请实施例提供的另一种通信装置的结构示意图。

具体实施方式

下面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、完整的描述。

首先，为了更好的理解本申请实施例公开的信道接入方法，对本申请实施例适用的通信系统进行描述。

一、通信系统。

请参见图 1，图 1 为本申请实施例提供的一种通信系统的结构示意图。该通信系统可包括但不限于一个接入点 (access point, AP)、两个站点 (station, STA)。图 1 所示的设备数量和形态用于举例，并不构成对本申请实施例的限定，实际应用中可以包括两个或两个以上的 AP，两个以上的 STA。图 1 所示的通信系统以 AP 101，STA 1021 和 STA 1022，且该 AP 101 能够为 STA 1021、STA 1022 提供无线服务为例进行阐述。其中，图 1 中的 AP 101 以基站为例，STA 1021 和 STA 1022 以手机为例。

本申请实施例中，上述通信系统可以为无线局域网 (Wireless local area network, WLAN) 或蜂窝网，或其他支持多条链路并行进行传输的无线通信系统。本申请实施例主要以部署 IEEE 802.11 的网络为例进行说明，而本申请涉及的各个方面可以扩展到采用各种标准或协议的其它网络，例如，BLUETOOTH (蓝牙)，高性能无线 LAN (high performance radio LAN, HIPERLAN) (一种与 IEEE 802.11 标准类似的无线标准，主要在欧洲使用) 以及广域网 (WAN)、个人区域网 (personal area network, PAN) 或其它现在已知或以后发展起来的网络。因此，

无论使用的覆盖范围和无线接入协议如何，本申请提供的各种方面可以适用于任何合适的无线网络。

本申请实施例中，STA 具有无线收发功能，可以支持 802.11 系列协议，与 AP 或其他 STA 进行通信。例如，STA 可以是允许用户与 AP 通信进而与 WLAN 通信的任何用户通信设备，如包括但不限于，平板电脑、桌面型、膝上型、笔记本电脑、超级移动个人计算机 (Ultra-mobile Personal Computer, UMPC)、手持计算机、上网本、个人数字助理 (Personal Digital Assistant, PDA)、手机等可以联网的用户设备，或物联网中的物联网节点，或车联网中的车载通信装置等。可选的，STA 还可以为上述这些终端中的芯片和处理系统。

本申请实施例中，AP 是为 STA 提供服务的装置，可以支持 802.11 系列协议。例如，AP 可以为通信服务器、路由器、交换机、网桥等通信实体，或，AP 可以包括各种形式的宏基站，微基站，中继站等，当然 AP 还可以为这些各种形式的设备中的芯片和处理系统，从而实现本申请实施例的方法和功能。

为了便于理解本申请公开的实施例，作以下两点说明。

(1) 本申请公开的实施例中场景以无线通信网络中无线局域网 (Wireless Fidelity, Wi-Fi) 网络的场景为例进行说明，应当指出的是，本申请公开的实施例中的方案还可以应用于其他无线通信网络中，相应的名称也可以用其他无线通信网络中的对应功能的名称进行替代。

(2) 本申请公开的实施例将围绕包括多个设备、组件、模块等的系统来呈现本申请的各个方面、实施例或特征。应当理解和明白的是，各个系统可以包括另外的设备、组件、模块等，并且/或者可以并不包括结合附图讨论的所有设备、组件、模块等。此外，还可以使用这些方案的组合。

二、本申请所要解决的技术问题。

目前，通信系统采用载波侦听多路访问/冲突避免 (carrier sense multiple access / collision avoidance, CSMA/CA) 机制避免在共享信道上产生冲突。即如图 2 所示，具有侦听能力的 STA 1 (即 CSMA/CA node) 在报文 (packet) 到达时，采用随机退避机制进行信道接入，即侦听随机时长 (T_s) 内的信道状态。若该随机时长内信道处于空闲状态，则该 STA 接入信道，即发送报文 y (即 packet y)。但只有在同样具有侦听能力的 STA 2 侦听信道，并且 STA 2 侦听信道的的时间 T 不等于 T_s 时，STA 1 才不会与 STA 2 发生冲突，即 STA 1 才能成功发送报文。也就是说，若 STA 2 的侦听时间 T 等于 STA 1 的侦听时间，那么 STA 1 和 STA 2 均认为在侦听时间内，信道空闲，均确定接入信道，即 STA 1 和 STA 2 在相同时间上发送报文，STA 1 发送报文 x，STA 2 发送报文 y，会造成 STA 1 和 STA 2 在共享信道上的冲突，从而 STA 1 和 STA 2 均不能成功发送报文。

该 CSMA/CA 机制可被认为是一种冲突分解类算法，即期望通过完全随机化达到冲突分解的效果。也就是说，该方式中的每个 STA 不具有预测其他 STA 是否接入信道的能力，从而导致系统的吞吐量较低、时延较高。同时，随着网络中 STA 数量的增加，网络中碰撞增多，导致 STA 的平均退避时间增大，进而也会导致传输时延较大、时延抖动较大。另外，研究证明 CSMA/CA 的理论容量上界只有约 85%，也就是在最好的情况下各 STA 之间仍然存在 15% 的冲突。而且，STA 的配置参数也会对实际性能产生较大的影响，研究表明一般情况下系统的容量只有 70%-80%。也就是说，通信系统采用 CSMA/CA 机制解决各 STA 之间的冲突时，存在吞吐量较低的问题。

人工智能 (artificial intelligence, AI) 技术被广泛应用于无线通信领域以提高通信性能和用户体验。其中，强化学习 (reinforcement learning, RL) 是一种适合信道接入问题的 AI 技

术, 它通过智能体(网络节点)在环境(无线网络)中采取动作(传输或不传输)的探索过程进行学习, 以找到最优策略来最大化期望奖励(吞吐量)。RL 的在线学习和无模型优化的特点使其相比传统基于模型的优化方法具有更好的泛化能力。

本申请实施例将 RL 技术与信道接入结合, AP 采用强化学习方法根据 N 个 STA 上报的动作信息训练每个 STA 的神经网络, 以获得每个 STA 对应神经网络的训练结果, 从而可使得每个 STA 根据该训练结果确定是否接入信道, 进而有利于提高 STA 预测是否接入信道的能力。

三. 信道接入方法 100 (每个 STA 向 AP 上报动作信息)。

本申请实施例提供一种信道接入方法 100。图 3 是该信道接入方法 100 的交互示意图。该信道接入方法 100 从 AP 与 STA 之间交互的角度进行阐述。该信道接入方法 100 包括但不限于以下步骤:

S101. N 个站点 STA 向接入点 AP 分别上报动作信息, N 个动作信息用于确定每个 STA 的第一神经网络的训练结果, N 为正整数。

其中, AP 对应有 M 个 STA, M 为大于 N 的正整数。N 个 STA 是 M 个 STA 中成功向 AP 上报动作信息的 STA。例如, 通信系统中的 AP#1 对应有 10 个 STA, 10 个 STA 中的 8 个 STA 成功向 AP 上报了动作信息, 也即 AP#1 接收到了 10 个 STA 中的 8 个 STA 上报的动作信息, 那么 N 等于 8。

针对 N 个 STA, 每个 STA 向 AP 上报一个动作信息, 所以 N 个 STA 上报的为 N 个动作信息。该动作信息是指一段时间的动作, 该动作为发送或不发送。其中, 一段时间包括多个时隙。该多个时隙是 STA 在上一次成功上报动作信息的时刻距离当前时刻内的多个时隙。例如, STA 1 上一次成功上报动作信息是在 t_0 时刻成功上报的, 当前时刻为 t_1 时刻, 那么多个时隙是指 t_0 - t_1 之间的多个时隙。也就是说, 每个 STA 上报的动作信息均包括多个时隙上的动作。每个 STA 上报的动作信息可用 a_i^t 表示, t 和 i 均为正整数, a_i^t 代表 STA i 在第 t 个时隙上的动作信息, 也即 STA i 在第 t 个时隙上是否发送了报文。

另外, 上述动作信息携带于 STA 上报的第一帧中。可理解的, 每个 STA 采用自身的第一帧携带动作信息, 然后向 AP 上报该第一帧。第一帧包括类型(Category)字段和动作详细信息(Action Details)信息字段, 类型字段用于指示第一帧的类型, 动作详细信息字段用于指示 STA 上报的动作信息。

一种可选的实施方式中, 上述第一帧为 STA 新增加的一个管理帧。例如, STA 新增一个管理帧为帧 1, 该帧 1 用于携带动作信息。帧 1 的帧结构如图 4 所示, 帧 1 包括类型(Category)字段和动作详细信息(Action Details)字段, 该 Category 字段用于指示帧 1 的类型, 该 Action Details 字段用于指示动作信息, 且动作信息携带于训练数据单元(training data element)子字段中。

另一种可选的实施方式中, 上述第一帧为协议中已有的管理帧中的一种帧。比如, 第一帧为服务质量动作(Quality of Service Action, QoS Action)帧, 其帧结构如图 5 所示。此时 Category 字段指示的第一帧的类型为 QoS Action 帧, Action Details 字段中的 QoS Action 子字段紧跟在 Category 字段后。STA 采用 QoS Action 字段中未被使用的值来指示上报动作信息, 即指示 Action Details 字段中的 training data element 子字段的内容。例如, QoS Action 字段包括两个比特(bite), 其两个 bite 代表的 00、01、11 值已被使用, 但 10 值未被使用, 那么 STA 采用该 10 值来指示上报的动作信息, 即采用该 10 值指示 training data element 的内容。

上述用于指示动作信息的 training data element 的单元格式可参见图 6 (a)。如图 6 (a)

所示, training data element 包括单元身份标识子字段 (element Identification, Element ID)、长度 (length) 子字段、单元身份标识扩展 (Element ID extension) 子字段以及训练数据 (Training data) 子字段。当目前的 Element ID 子字段中的数值全被使用时, element ID 子字段和 Element ID extension 子字段联合指示 Training data 的 ID。Length 子字段用于表示 Training data 的长度。Training data 代表的是 STA 上报的动作信息。

当每个 STA 的第一帧中 Training data 的单元格式不相同, STA 上报的动作信息的内容也不相同。以下结合 Training data 的单元格式, 阐述动作详细信息字段的几种可选实施方式, 也即阐述动作信息的可选实施方式。

1. 动作详细信息字段包括时间指示子字段、数据 1 子字段至数据子 T 子字段, T 为正整数。

Training data 的单元格式可参见上述图 6 (a) 所示, Training data 包括时间, 数据 1 至数据 T。那么, 动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段。

其中, 时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻, 时间指示子字段可以由时间戳、序列号等方式实现。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息, 比如第一响应信息是确认信息 (Acknowledge, ACK)。也就是说, STA 接收到第一响应信息, 表明 STA 成功上报了动作信息。因此, 时间指示子字段是用于指示 STA 上一次成功上报动作信息的时间。

数据 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个时隙内的动作, 即数据 1 子字段用于指示 STA 自上一次成功上报动作信息以来, 在第一个时隙内的动作。数据 T 子字段用于指示 STA 在上一次成功接收第一响应信息后的第 T 个时隙内的动作, 即数据 T 子字段用于指示 STA 自上一次成功上报动作信息以来, 在第 T 个时隙内的动作。

也就是说, 每个 STA 向 AP 上报动作信息时, 上报的是 STA 上一次成功上报动作信息的时间, 以及自上一次成功上报动作信息以来在每个时隙上的动作, 从而使得 AP 获得每个 STA 自上一次成功上报动作信息以来, 在每个 slot 上侦听的动作。

2. 动作详细信息字段包括时间指示子字段、动作 1 子字段至动作 P 子字段、.....、持续时间 1 子字段至持续时间 P 子字段, P 为正整数。

Training data 的单元格式可参见上述图 6 (b) 所示, 与图 6 (a) 不同的是, Training data 包括起始时间、动作 1 和时间 1、.....、动作 P 和时间 P。那么, 动作详细信息字段包括时间指示子字段、动作 1 子字段和时间 P 子字段、.....、动作 P 子字段至时间 P 子字段。

其中, 时间指示子字段用于指示 STA 上一次成功接收第一响应信息的接收时刻, 第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。那么, 时间指示子字段用于指示 STA 上一次成功上报动作信息的时刻。

动作 1 子字段用于指示 STA 上一次成功接收第一响应信息的后的第一个动作, 即动作 1 子字段用于指示 STA 在上一次成功上报动作信息后的第一个动作。时间 1 子字段用于指示动作 1 的持续时间或动作 1 的结束时间。动作 P 子字段用于指示 STA 在上一次成功接收第一响应信息的时刻距离当前时刻内的第 P 个动作, 即动作 P 子字段用于指示 STA 在上一次成功上报动作信息后距离当前时刻内的第 P 个动作。时间 P 子字段用于指示动作 P 的持续时间或动作 P 的结束时间。

可理解的, 动作 1 为 STA 在上一次成功上报动作信息后的第一个动作。当时间 1 子字段用于指示动作 1 的持续时间, 时间 P 子字段用于指示动作 P 的持续时间时, 动作 1 不发生变化时, 持续时间 1 不断累加; 当动作 1 发生改变时, 新增动作 2, 并记录动作 2 的持续时间 2, 直至记录到当前时刻前的最后一个动作 (即动作 P) 时, STA 向 AP 上报所记录的动作信息,

即向 AP 上报上一次成功上报动作信息的时刻、动作 1 和动作 1 的持续时间、动作 2 和动作 2 的持续时间，直至动作 P 和动作 P 的持续时间。

例如，STA 1 在上一次成功上报动作信息后的第 1 个 slot 内不发送报文，则记录动作 1 为不发送。STA 1 在第 1 个 slot 到第 3 个 slot 内均不发送报文，则持续时间 1 累加为 3 个 slot。STA 1 第 4 个时隙上不发送报文的动作变为发送报文，则 STA 1 增加动作 2，动作 2 为发送，若发送报文的动作持续到当前时刻（第 9 个 slot），则 STA 1 记录动作 2 的持续时间 2 为 6 个 slot。因此 STA 1 向 AP 上报的动作信息包括 STA 1 上一次成功上报动作信息的时间，动作 1 为不发送，且不发送的持续时间为 3 个 slot，动作 2 为发送，且发送的持续时间为 6 个 slot。

也就是说，每个 STA 上报的是自身上一次成功上报动作信息的时间、STA 在上一次成功上报动作信息的时刻距离当前时刻内所发生的多个动作，以及每个动作所持续的时间或每个动作的结束时间。该实施方式有利于 AP 获知每个 STA 自上一次成功上报动作信息以来，在每个时隙上的动作行为。

3. 动作信息字段包括时间 1 指示子字段和动作 1 子字段、.....、时间 P 指示子字段和动作 P 子字段，P 为正整数。

Training data 的单元格式可参见上述图 6(c) 所示，与图 6(a)、图 6(b) 不同的是 Training data 包括时间 1 和动作 1、时间 2 和动作 2、.....、时间 P 和动作 P。那么，动作详细信息字段包括时间 1 指示子字段和动作 1 子字段、.....、时间 P 指示子字段和动作 P 子字段。

其中，时间 1 指示子字段用于指示动作 1 的起始时间。动作 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个动作，第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息，那么动作 1 子字段用于指示 STA 在上一次成功上报动作信息后的第一动作。时间 P 指示子字段用于指示动作 P 的起始时间。动作 P 子字段用于指示 STA 在上一次成功接收第一响应信息的时刻距离当前时刻内的第 P 个动作，即动作 P 子字段用于指示 STA 在上一次成功发送动作信息的时刻距离当前时刻内的第 P 个动作。

可理解的，动作 1 为 STA 在上一次成功上报动作信息后的第一个动作，时间 1 标记动作 1 的起始时间。若动作 1 发生改变，则 STA 记录动作 2 和动作 2 的起始时间（时间 2），直至记录到上一次成功上报动作信息的时刻距离当前时刻内多个动作中的最后一个动作和该动作的起始时间（动作 P 和时间 P）时，STA 向 AP 上报所记录的动作信息。

也就是说，每个 STA 向 AP 上报的是自身自上一次成功上报动作信息以来，所发生的每个动作和每个动作的起始时间，从而有利于 AP 根据每个 STA 上报的动作和每个动作的起始时间，获得每个 STA 在多个时隙上的发送或不发送报文的行为信息。

4. 动作信息字段包括时间 1 指示子字段和持续时间 1 子字段、.....、时间 K 指示子字段和持续时间 K 子字段，K 为正整数。

Training data 的单元格式可参见上述图 6(d) 所示，与图 6(a)-图 6(c) 不同的是 Training data 包括时间 1 和持续时间 1、时间 2 和持续时间 2、.....、时间 K 和持续时间 K。那么，动作详细信息字段包括时间 1 指示子字段和持续时间 1 子字段、.....、时间 K 指示子字段和持续时间 K 子字段。

其中，时间 1 指示子字段用于指示动作 1 的开始时间/结束时间。动作 1 是 STA 在上一次成功接收第一响应信息后，第一次发送报文且未接收到第二响应信息时的发送动作。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息，第二响应信息是 AP 成功接收 STA 发送的报文时发送的响应信息。那么，动作 1 是 STA 在上次一成功上报动作信息后，第一次发送报文且未发送成功时的动作。持续时间 1 子字段用于指示动作 1 的持续时间，即

持续时间 1 子字段指示的是动作 1 发送的报文的包长。

时间 K 指示子字段用于指示动作 K 的开始时间/结束时间。动作 K 是 STA 在上一次成功接收第一响应信息后，第 K 次发送报文且未接收到第二响应信息的发送动作。那么，动作 K 是 STA 在上一次成功上报动作信息后，第 K 次发送报文且未成功发送时的动作。持续时间 K 子字段用于指示动作 K 的持续时间，即持续时间 K 子字段指示的是动作 K 发送的报文的包长。

这是由于 AP 只有在多个 STA 同时发送报文，信道发生冲突时，无法获知哪些 STA 尝试接入信道。因此，每个 STA 只需在发送报文不成功时，向 AP 上报该动作信息，即每个 STA 上报的是每个发送报文不成功时的发送动作，以及该动作的开始时间/结束时间、每次发送的报文的包长，以使得 AP 获知在信道发生冲突时，哪些 STA 在尝试接入信道。

5. 动作信息字段包括第一时间 1 指示子字段和第二时间 1 指示子字段、.....、第一时间 K 指示子字段和第二时间 K 指示子字段，K 为正整数。

Training data 的单元格式可参见上述图 6(e) 所示，与图 6(a)-图 6(d) 不同的是 Training data 包括第一时间 1 和第二时间 1、.....、第一时间 K 和第二时间 K。那么，动作详细信息字段包括第一时间 1 指示子字段和第二时间 1 指示子字段、.....、第一时间 K 指示子字段和第二时间 K 指示子字段。

其中，第一时间 1 子字段用于指示动作 1 的起始时间。动作 1 是 STA 在上一次成功接收第一响应信息后，第一次发送报文且未接收到第二响应信息时的发送动作。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息，第二响应信息是 AP 成功接收 STA 发送的报文时发送的响应信息。那么，动作 1 是 STA 在上次一成功上报动作信息后，第一次发送报文且未发送成功时的动作。第二时间 1 指示子字段用于指示动作 1 的结束时间。

第一时间 K 子字段用于指示动作 K 的起始时间。动作 K 是 STA 在上一次成功接收第一响应信息后，第 K 次发送报文且未接收到第二响应信息的发送动作。那么，动作 K 是该 STA 在上一次成功上报动作信息后，第 K 次发送报文且未成功发送时的动作。第二时间 K 用于指示动作 P 的结束时间。

可见，动作 1 至动作 K 均是 STA 在上一次成功上报动作信息后，发送报文不成功时的动作。该情况下，每个 STA 向 AP 上报的是在上一次成功上报动作信息后，每一次发送报文不成功时的开始时间和结束时间，从而有利于 AP 根据每一次发送报文不成功时的开始时间和结束时间，确定每次发送报文不成功是在哪个时隙上发送的，以及发送的报文的包长，以进一步获得每个 STA 在每个时隙上的行为信息。

可见，上述 5 种 Training data 字段的不同格式单元，代表了每个 STA 上报的动作信息中的不同内容，从而 STA 向 AP 上报的动作信息更加灵活。

可理解的，每个 STA 向 AP 上报动作信息的时间是 AP 预定义的。例如，AP 预定义每个 STA 按照预设周期，向 AP 上报动作信息，那么每个 STA 每间隔该预设周期就向 AP 上报动作信息。另外，AP 给每个 STA 预定义的上报时间可以是不相同的。例如，AP 给 STA 1 预定义每间隔预设时间 1，向 AP 上报动作信息，而给 STA 2 预定义每间隔预设时间 2，向 AP 上报动作信息。

可选的，每个 STA 向 AP 上报动作信息的时间是 AP 通过信令告知给每个 STA 的。例如，AP 通过下行控制信息 (downlink control information, DCI) 向每个 STA 告知上报动作信息的时间。再例如，AP 通过 DCI#1 向 STA 1 告知 STA 1 上报动作信息的时间#1，通过 DCI#2 向 STA 2 告知上报动作信息的时间#2。

S102.AP 接收 N 个 STA 分别上报的动作信息。

S103.AP 根据 N 个动作信息，确定每个 STA 的第一神经网络的训练结果。

可理解的，AP 根据 N 个动作信息，训练每个 STA 的第一神经网络，获得每个 STA 的第一神经网络的训练结果。例如，5 个 STA 一共上报了 5 个动作信息，5 个 STA 分别对应第一神经网络#1-第一神经网络#5，AP 根据该 5 个动作信息训练 STA 1 的第一神经网络#1，获得第一神经网络#1 的训练结果，根据该 5 个动作信息训练 STA 2 的第一神经网络#2，获得第一神经网络#2 的训练结果，直至获得 STA 5 的第一神经网络#5 的训练结果。

可理解的，第一神经网络的训练结果为第一神经网络的神经网络参数或梯度。该神经网络参数为第一神经网络中神经元的权重和偏置。例如，第一神经网络的结构如图 7 所示，包含一个输入层，一个输出层，及多个中间层，且每层包括多个节点，该节点称为神经元。其中，相邻两层的神经元间两两相连。

对于相邻两层的神经元而言，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。用矩阵可以表示为：

$$h = f(wx + b) \quad (1)$$

其中 w 为权重矩阵， b 为偏置向量， f 为激活函数。则 n 层神经网络的输出 y 可以递归表达为：

$$y = f_n(w_n f_{n-1}(\dots) + b_n) \quad (2)$$

也就是说，可将第一神经网络理解为一个从输入 x 到输出 y 的映射关系。神经网络的训练过程是指从已有数据得到该映射关系的过程，即得到 w 和 b 的过程。第一神经网络的训练结果可以是神经网络参数 w 和 b 。

另外，AP 可采用梯度下降法对神经网络进行训练，因此，神经网络的训练结果也可以是梯度。梯度是神经网络的损失函数对该神经网络参数的偏导，即神经网络的损失函数对上述 w 和 b 的偏导。

神经网络参数/梯度均是用于对应 STA 更新对应的第一神经网络，即哪个 STA 的神经网络参数/梯度就是用于更新哪个 STA 的第一神经网络。例如，神经网络参数#1 是 STA 1 的神经网络参数，那么神经网络参数#1 是用于 STA 1 更新 STA 1 的第一神经网络。

一种可选的实施方式中，AP 根据 N 个动作信息，确定每个 STA 的第一神经网络的训练结果，是指：AP 将每个 STA 的状态信息输入到对应 STA 的第一神经网络，获得第一神经网络的输出值；AP 将每个第一神经网络的输出值输入到第二神经网络，获得第二神经网络的输出值，第二神经网络的输出值用于表征预设时间内的期望奖励；AP 根据第二神经网络的输出值和奖励函数训练第三神经网络，通过最小化第三神经网络的损失函数，确定每个第一神经网络的训练结果，第三神经网络包括每个第一神经网络和第二神经网络。

其中，STA 的状态信息是根据 STA 的动作信息获得的，第二神经网络的神经网络参数是根据 N 个动作信息获得的，奖励函数是根据 N 个动作信息确定的。

可理解的，AP 获得每个 STA 上报的动作信息后，根据每个的动作信息确定载波侦听结果信息或报文传输结果信息，再根据 N 个动作信息和 N 个载波侦听结果信息确定状态信息，或者根据 N 个动作信息和 N 个报文传输结果信息确定状态信息。载波侦听结果信息或报文传输结果信息包括 o_t^i 、 l_t^i 、 d_t^i 、 d_t^{-i} 。

上述 $a_t^i=0$ 时，表示不发送；当时 $a_t^i=1$ ，表示发送。 o_t^i 代表载波侦听结果或报文传输结果。 o_t^i 在 $a_t^i=0$ 时， o_t^i 代表载波侦听结果，即表示 STA i 在第 t 个时隙上的载波侦听结果， $o_t^i=0$ 表示信道空闲， $o_t^i=1$ 表示信道忙碌。 o_t^i 在 $a_t^i=1$ 时， o_t^i 代表报文传输结果信息，即表示 STA i 在第 t 个时隙上的发送报文的结果， $o_t^i=0$ 表示报文发送成功， $o_t^i=1$ 表示报文发送失败。也就是

说, o_t^i 可 a_t^i 根据的值, 代表不同的含义, 在 $a_t^i=0$ 时, 代表载波侦听结果, 在 $a_t^i=1$ 时, 代表报文传输结果。

l_t^i 表示 a_t^i 和 o_t^i 的持续时间。 d_t^i 表示 STA i 上一个第二响应信息的接收时刻距离第 t 个时隙的时长。 d_t^{-i} 表示 STA j 在上一个第二响应信息的接收时刻距离第 t 个时隙的时长, STA j 为除 STA i 外的其他任一 STA。

如图 8 所示, d_t^i 表示 STA i 上一个第二响应信息的接收时刻距离第 t 个时隙的时长, 即 d_t^i 表示的是当前时刻距离 STA i 在上一次成功发送报文的时长。 d_t^{-i} 表示 STA j 在上一个第二响应信息的接收时刻距离第 t 个时隙的时长, STA j 为除 STA i 外的其他任一 STA, 即 d_t^{-i} 表示的是除 STA i 外的任一 STA 在当前时刻距离上一次成功发送报文的时长。该 d_t^i 和 d_t^{-i} 可以是 STA i 通过侦听信道上的响应得到的。STA i 在侦听到本 STA 发送报文成功的响应时, $d_t^i=0$, 在未侦听到本 STA 发送报文成功的响应时, $d_t^i=d_t^i+1$, 即当前时刻距离 STA i 在上一次成功发送报文的时长继续叠加。STA i 侦听到信道对其他 STA 的响应时, $d_t^{-i}=0$, 在未侦听到信道对其他 STA 的响应时, $d_t^{-i}=d_t^{-i}+1$, 即除 STA i 外的任一 STA 在当前时刻距离上一次成功发送报文的时长继续叠加。

可见, AP 是先将根据每个 STA 上报的信息获得的状态信息输入到该 STA 的第一神经网络中, 获得每个第一神经网络的输出值, 再将 N 个第一神经网络的输出值输入到第二神经网络中, 获得第二神经网络的输出值, 然后根据损失函数训练第三神经网络, 最终获得第一神经网络的训练结果。每个 STA 的第一神经网络的训练结果是根据 N 个 STA 上报的信息确定的, 而不是只根据自身 STA 的信息确定的, 从而有利于提高每个 STA 对其他 STA 信道接入行为的预测能力。

以下以 AP 采用目标 Q 神经网络训练每个第一神经网络为例阐述 AP 的训练过程。

目标 Q 网络的训练示意图如图 9 所示。图 9 中包括目标 Q 网络 (target Q network) 和预测 Q 网络 (prediction Q network)。目标 Q 网络和预测 Q 网络的结构如图 10 所示。图 10 所示的神经网络包括代理网络 1 (agent network 1) 至代理网络 N (agent network N), 以及混合网络 (Mixing network)。Agent network 1 至 agent network N 是 STA 1 至 STA N 的第一神经网络, 即每个 agent network 对应一个 STA。Mixing network 为上述的第二神经网络。

每个 agent network 的输入为对应 STA 过去一段时间的状态信息, 即 $\tau_t^i, \tau_t^i = [c_{t-T+1}^i, \dots, c_{t-1}^i, c_t^i]$, τ_t^i 表示 STA i 在当前时刻 t 前的连续 T 个时刻的状态信息, 每个 STA 的状态信息是根据该 STA 上报的动作信息获得的。STA 上报的动作信息为 a_t^i , AP 根据该 STA 的 a_t^i , 获得 o_t^i, l_t^i, d_t^i 和 d_t^{-i} , 从而 STA 根据 $a_t^i, o_t^i, l_t^i, d_t^i$ 和 d_t^{-i} 获得自身 STA 的状态信息, 即 $c_{t+1}^i = (a_t^i, o_t^i, l_t^i, d_t^i, d_t^{-i})$ 。首先, AP 将每个 STA 的状态信息输入到对应的 agent network, 获得该 agent network 的输出值, agent network 的输出值为 $Q_i(\tau_t^i, a_t^i)$ 。每个 agent network 包括门控循环单元 (Gated Recurrent Unit, GRU), 全连接神经网络 (fully connected neural network, FC) 和模块 π 。每个 STA 的状态信息经过 GRU 和 FC 后, 模块 π 再根据 ϵ -greedy 算法进行动作选择, 进而确定 $Q_i(\tau_t^i, a_t^i) \circ Q_i(\tau_t^i, a_t^i)$ 表示 STA i 在预设时间内的期望奖励。然后, AP 将每个 agent network 的输出值作为 mixing network 的输入, 获得第二神经网络的输出值, 第二神经网络的输出值为 $Q_{tot}(\tau_t, a_t, s_t)$ 。mixing network 用于将 N 个 agent network 的输出值进行汇总处理。AP 获得的第二神经网络的输出值表示整个网络在预设时间内的期望奖励。此外, 第二神经网络的神经网络参数是根据 N 个动作信息获得的, 即第二神经网络的神经网络参数是由 N 个动作信息决定的。图 10 中的 GRU 32 表示该 GRU 的隐状态具有 32 个神经元, FC 32 表示该 FC 具有 32 个神经元, 同理 FC 2 表示该 FC 具有 2 个神经元。

AP 根据 mixing network 的输出值和奖励函数计算第三神经网络的损失函数, 并通过最小化该损失函数训练第三神经网络, 即训练每个 agent network 和 mixing network, 进而确定每个 agent network 的神经网络参数。第三神经网络的损失函数为:

$$L(\theta) = \frac{1}{|E|} \sum_{e(t) \in E} [r(t) + \gamma \max_{\theta^-} Q_{tot}(\tau_{t+1}, a_{t+1}, s_{t+1}; \theta^-) - Q_{tot}(\tau_t, a_t, s_t; \theta)]^2 \quad (3)$$

其中, $r(t)$ 表示奖励函数, γ 示折扣因子 (discount factor), 通常 $\gamma=0.9$, e_t 表示经验 (experience), E 表示经验池, $|E|$ 表示经验池中经验 e_t 的数量, $e(t) = (s_t, \tau_t, a_t, r_t, s_{t+1}, \tau_{t+1})$, $\tau_t = [\tau_t^1, \tau_t^2, \dots, \tau_t^N]$, $a_t = [a_t^1, a_t^2, \dots, a_t^N]$, $s(t) = [d_t^1, d_t^2, \dots, d_t^N, a_t^1, a_t^2, \dots, a_t^N]$, $Q_{tot}(\tau_{t+1}, a_{t+1}, s_{t+1}; \theta^-)$ 表示图 8 中目标 Q 网络的输出, θ^- 为目标 Q 网络的神经网络参数, $Q_{tot}(\tau_t, a_t, s_t; \theta)$ 表示图 8 中预测 Q 网络的输出, θ 为预测 Q 网络的神经网络参数。mixing network 的神经网络参数由 $s(t)$ 决定。

AP 训练第三神经网络的过程可参见上述图 9 所示的示意图。即 AP 利用小批量梯度下降的方法更新 Q 网络的神经网络参数。可理解的, AP 每次将 θ^- 的值固定, 然后利用损失函数和 mixing network 的输出值训练预测神经网络的神经网络参数 θ , 每训练完 C 次, 将神经网络参数 θ 作为目标神经网络的固定参数 θ^- , 再迭代训练预测 Q 网络的神经网络参数, 通过最小化第三神经网络的损失函数, 确定每个 agent network 的训练数据, 通常 $C=100$ 。

关于上述第三神经网络的奖励函数的计算有以下几种可选的实施方式:

1. 将第三神经网络的奖励函数设置为 1。

可理解的, AP 在根据动作信息确定第一 STA 发送报文成功时, 将第三神经网络的奖励函数设置为 1。第一 STA 是上述 N 个 STA 中, 上一次成功接收第二响应信息的时刻距离当前时刻的时间间隔最长的 STA, 即第一 STA 是距离上一次成功发送报文时长最长的 STA。

也就是说, 当 AP 根据 N 个动作信息确定距离上一次成功发送报文时长最长的 STA 在多个时隙中发送报文成功, 则将奖励函数的值设置为 1。即 $r_t = 1$, $i = \arg \max [d_t^1, d_t^2, \dots, d_t^N]$, d_t^i 表示 STA i 上一个第二响应信息的接收时刻距离第 t 个时隙的时长, $\arg \max$ 表示取各数值中数值最大的值对应的 i。

2. 将奖励函数的值设置为第一时长减 1。

可理解, AP 在根据 N 个动作信息确定第二 STA 发送报文成功时, 将奖励函数的值设置为第一时长减 1, 即 $r_t = d_t^j - 1$, $j = \arg \max [d_t^1, d_t^2, \dots, d_t^N]$ 。第二 STA 是 N 个 STA 中, 除第一 STA 外的 STA, 第一 STA 是 N 个 STA 中上一次成功接收第二响应信息的时刻距离当前时刻的时间间隔最长的 STA。第一时长是第二 STA 上一次成功接收第二响应信息的时刻距离当前时刻的时长。也就是说, 当 AP 根据 N 个动作信息确定除上一次成功发送报文的时长最长的 STA 外的 STA 在多个时隙中的某个时隙上发送报文成功时, 将奖励函数的值设置为第一时长减 1。

3. 将奖励函数的值设置为 -1。

可理解的, AP 在根据 N 个动作信息确定 N 个 STA 中的 M 个 STA 在同一时隙上发送报文时, 将奖励函数的值设置为 -1, 即 $r_t = -1$ 。M 为小于或等于 N 的正整数。也就是说, AP 根据 N 个动作信息确定 N 个 STA 中的部分 STA 在统一时隙上发送报文时, 表明该部分 STA 在该时隙上发送报文发生了信道冲突, 该部分 STA 均不能成功发送报文, 即将奖励函数的值累减, 即将奖励函数的值减 1。

4. 将奖励函数的值设置为 0。

可理解的, AP 在根据 N 个动作信息确定 N 个 STA 在一个时隙上均未发送报文时, 将奖励函数的值设置为 0, 即 $r_t = 0$ 。也就是说, AP 在根据 N 个动作信息确定所有的 STA 均未在一个时隙上发送报文时, 对于未来的期望奖励没有, 因此将奖励函数的值设置为 1。

另外, 除上述 4 中情况外, AP 也可将奖励函数的值设置为 0。

本申请实施例中, 若每个 STA 上报动作信息的时间不相同, 或者 N 个 STA 中的部分 STA 上报动作信息的时间不相同, 那么 AP 在当前时刻训练神经网络时, 可能某些 STA 并未上报

动作信息, 只有部分 STA 上报了最新的动作信息。该情况下, AP 在训练每个 STA 的神经网络时, 使用当前时刻上报的动作信息, 以及当前时刻未上报动作信息的 STA 的上一次上报的动作信息联合训练每个 STA 的第一神经网络, 从而也可实现对每个 STA 的第一神经网络的集中训练。另外, 该方式可使得在当前时刻其动作信息未发生变化的 STA 不需上报动作信息, 从而可减少通信系统的信令开销。

与目前的 STA 根据自身观测的发送行为、发包时长训练本 STA 的神经网络相比, 本申请实施例中, AP 是根据 N 个 STA 的 N 个动作信息训练每个 STA 的第一神经网络的, 即训练每个 STA 的第一神经网络时, 参考了 N 个 STA 的动作信息, 从而可使得 AP 对每个第一神经网络的训练更优, 能获得更好的训练结果, 使得第一神经网络的预测能力更好。

S104. AP 将每个 STA 的第一神经网络的训练结果发送给对应的 STA。

S105. 针对每个 STA, STA 接收来自 AP 的第一神经网络的训练结果。

S106. 针对每个 STA, STA 根据第一神经网络的训练结果, 更新第一神经网络, 并在侦听到信道空闲时根据更新后的第一神经网络和该 STA 当前的状态信息确定是否接入信道。

其中, 该 STA 当前的状态信息包括自身过去一段时间的动作、载波侦听结果, 以及报文传输结果。

一种可选的实施方式中, 如上所述, 第一神经网络的训练结果为第一神经网络的神经网络参数。那么 STA 根据第一神经网络的训练结果, 更新第一神经网络, 是指 STA 将第一神经网络之前的神经网络参数, 更新为接收到的神经网络参数, 获得更新后的第一神经网络。

另一种可选的实施方式中, 如上所述, 第一神经网络的训练结果为第一神经网络的梯度。那么 STA 根据第一神经网络的训练结果, 更新第一神经网络, 是指 STA 将梯度进行计算处理, 获得第一神经网络的神经网络参数, 再将第一神经网络原来的神经网络参数替换为该神经网络参数, 从而获得更新后的神经网络参数。STA 对梯度进行计算处理的过程为 $\theta' = \theta + \gamma g$, θ' 为更新后第一神经网络的神经网络参数, θ 更新前第一神经网络的神经网络参数。 γ 为第一神经网络的学习效率, g 为梯度。

一种可选的实施方式中, STA 根据第一神经网络的训练结果更新所述第一神经网络, 并在侦听到信道空闲时根据更新后的第一神经网络和侦听的动作信息确定是否接入信道, 包括: STA 将动作信息输入到更新后的第一神经网络输出第一值和第二值, 第一值用于表征接入信道获得的预期奖励, 第二值用于表征不接入信道获得的预期奖励; STA 在第一值大于第二值时, 确定接入信道; STA 在第一值小于第二值时, 确定不接入信道, 即 STA 在侦听到信道空闲时, 根据更新后的第一神经网络输出的第一值和第二值的大小, 确定是否接入信道。

以 STA 的第一神经网络为 Q 神经网络中的一部分为例, 阐述 STA 在侦听到信道空闲时, 根据第一神经网络的训练结果和当前时刻侦听到的动作信息确定是否接入信道的实施方式。此时, STA 的第一神经网络的结构如图 10 所示。STA 将当前时刻自身侦听信道获得的动作信息作为 agent network 的输入, 获得 $Q(a_i^t, 0)$ 和 $Q(a_i^t, 1)$, 即上述的第一值和第二值。当 $Q(a_i^t, 0)$ 大于 $Q(a_i^t, 1)$ 时, 表明该 STA 接入信道获得的期望奖励更大一些, 那么该 STA 确定接入信道; 当 $Q(a_i^t, 0)$ 小于 $Q(a_i^t, 1)$ 时, 表明该 STA 不接入信道获得的期望奖励更大一下, 那么该 STA 确定不接入信道。

本申请实施例中, STA 可在侦听到信道空闲时, 根据 AP 训练的第一神经网络的训练结果, 以及自身当前时刻侦听到的动作信息, 自行决定是否接入信道。第一神经网络的训练结果又是 AP 根据每个 STA 的动作信息对第一神经网络训练获得的, 第一神经网络的预测性较高, 因此该方式有利于 STA 在决定接入信道时, 能成功发送报文的概率较大, 即存在信道冲突的可能性较小, 进而可提高系统的吞吐量, 降低通信系统的时延。

本申请实施例的实现框图可参见图 11。图 11 的实现框图中包括 AP 的中心训练部分和 STA 的分布执行部分。AP 的中心训练部分和 STA 的分布执行部分均包括每个 STA 对应的第一神经网络，第一神经网络的神经网络参数为 θ_1 。

AP 的中心训练是指：AP 根据 N 个 STA 上报的 N 个动作信息获得的 N 个状态信息训练每个第一神经网络，获得每个第一神经网络的训练结果，即每个第一神经网络的训练结果均是根据 N 个动作信息获得，可提高第一神经网络的预测性。每个动作信息是每个 STA 对历史环境进行观测得到的。

每个 STA 的分布执行是指：每个 STA 获得 AP 下发的第一神经网络的训练结果后，该 STA 使用训练结果更新自身的第一神经网络，再在侦听到信道空闲时，根据侦听的动作信息和更新后的第一神经网络，自行决定是否接入信道。STA 根据更新后的第一神经网络确定是否接入信道的方式，可使得 STA 决定是否接入信道的预判性更高，进而可提高系统的吞吐量和降低系统的通信时延。

可理解的，本申请实施例适用于所有中心训练分布执行的多智能体强化学习算法。比如，Aho-Corasick 自动机算法、近端策略优化（Proximal Policy Optimization, PPO）算法、多智能体深度确定性策略梯度（Multi-Agent Deep Deterministic Policy Gradient, MADDPG）算法等。

本申请实施例中，N 个 STA 向 AP 上报动作信息，AP 根据 N 个 STA 上报的 N 个动作信息，确定每个 STA 的第一神经网络的训练结果，并将每个 STA 的第一神经网络的训练结果发送给对应的 STA，从而每个 STA 可根据第一神经网络的训练结果更新第一神经网络，并在侦听到信道空闲时根据更新后的第一神经网络和侦听动作信息确定是否接入信道。AP 是根据 N 个动作信息训练每个 STA 的第一神经网络的，可使得第一神经网络的预测性更好，从而有利于提高每个 STA 对其他 STA 信道接入行为的预测能力，即每个 STA 发送报文时，其信道冲突的概率更小，进而有利于提高系统的吞吐量和降低通信时延。

另外，本申请实施例与目前的研究中，STA 自行根据网络中所有的 STA 的历史动作信息训练自身的神经网络的实施方式相比，每个 STA 是根据 AP 下发的第一神经网络的训练结果和自身侦听到的历史动作信息，独立决策是否进行信道的接入，不依赖除本 STA 外其他 STA 的动作信息，因此对于每个 STA 而言，其实际操作性较好。

目前的研究中，每个 STA 还可分别自行训练自身的神经网络，并向 AP 上报训练获得的神经网络参数，然后 AP 将所有 STA 的神经网络参数进行处理后，获得新的神经网络参数，并将新的神经网络参数广播给每个 STA，再由 STA 根据新的神经网络参数确定是否接入信道。本申请实施例与该研究相比，每个 STA 的神经网络是由 AP 集中训练的，网络中的每个 STA 无需自行训练自身的神经网络，即无需网络中的每个 STA 具有独立训练神经网络的能力，也可降低每个 STA 与 AP 的交互，可减少系统的信令开销与计算能力。

图 12 为本申请实施例下的系统吞吐量与采用 CSMA/CA 技术解决信道冲突时的系统的吞吐量对比图。本申请实施例下的系统吞吐量明显高于采用 CSMA/CA 技术解决信道冲突时的吞吐量。图 13 为本申请实施例下系统的平均时延与采用 CSMA/CA 技术解决信道冲突时系统的平均时延对比图。本申请实施例下的平均时延明显低于采用 CSMA/CA 技术解决信道冲突时的平均时延。图 14 为本申请实施例下系统的时延抖动与采用 CSMA/CA 技术解决信道冲突时系统的时延抖动对比图。本申请实施例下的时延抖动明显低于采用 CSMA/CA 技术解决信道冲突时的时延抖动。

四、每个 STA 上报动作信息和载波侦听结果信息，或者每个 STA 上报动作信息和报文传

输结果信息。

可理解的，每个 STA 除了上报动作信息外，还可上报载波侦听结果信息或报文传输结果信息，其实施方式如下：

1. 每个 STA 上报动作信息和载波侦听结果信息。

也就是说，每个 STA 除了上报上述的动作信息外，还会上报载波侦听结果信息。载波侦听结果信息包括 STA 自上一次成功上报动作信息后，距离当前时刻内的多个时隙内每个时隙的载波侦听结果。那么，AP 接收 N 个 STA 分别上报的动作信息和载波侦听结果信息。

此时，N 个动作信息和 N 个载波侦听结果信息携带于 STA 上报的第一帧的动作详细信息字段中，第一帧的帧结构如上述图 5 所示，不再赘述。动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段，T 为正整数。时间指示子字段用于指示 STA 上一次成功接收第一响应信息时刻。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。那么时间指示子字段用于指示 STA 上一次成功上报动作信息的时刻。

数据 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个时隙内的动作和载波侦听结果。数据 T 子字段用于指示 STA 上一次成功接收第一响应信息后的第 T 个时隙内的动作和载波侦听结果。可理解的，数据 1 子字段用于指示 STA 自上一次成功上报动作信息以来，在第一个时隙内的动作和载波侦听结果。数据 T 子字段用于指示 STA 自上一次成功上报动作信息以来，在第 T 个时隙内的动作和载波侦听结果。

那么，上述 S103 中，AP 根据 N 个动作信息，确定每个 STA 的第一神经网络的训练结果，可指：AP 根据 N 个动作信息和 N 个载波侦听结果信息，确定每个 STA 的第一神经网络的训练结果。可理解的，AP 无需根据动作信息确定载波侦听结果信息，可直接使用接收到的动作信息和载波侦听结果信息确定 STA 的第一神经网络的训练结果，从而可减少 AP 处理的复杂度。

可选的，上述 S103 中，AP 根据 N 个动作信息，确定每个 STA 的第一神经网络的训练结果，还可指：AP 根据 N 个动作信息，以及根据 N 个动作信息确定的 N 个载波侦听结果信息，确定每个 STA 的第一神经网络的训练结果。也就是说，该实施方式下，即使 STA 上报了载波侦听结果信息，AP 仍可采用根据动作信息确定的载波侦听结果信息确定第一神经网络的训练结果。

2. 每个 STA 上报动作信息和报文传输结果信息。

也就是说，每个 STA 除了上报上述的动作信息外，还会上报报文传输结果信息。报文传输结果信息包括 STA 自上一次成功上报动作信息后，距离当前时刻内的多个时隙内 STA 发送报文时的报文传输结果。那么，AP 接收 N 个 STA 分别上报的动作信息和载波侦听结果信息。

此时，N 个动作信息和 N 个报文传输结果信息携带于 STA 上报的第一帧的动作详细信息字段中，第一帧的帧结构如上述图 5 所示，不再赘述。动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段，T 为正整数。时间指示子字段用于指示 STA 上一次成功接收第一响应信息时刻。第一响应信息是 AP 成功接收 STA 发送的动作信息时发送的响应信息。那么时间指示子字段用于指示 STA 上一次成功上报动作信息的时刻。

数据 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个时隙内的动作和报文传输结果。数据 T 子字段用于指示 STA 上一次成功接收第一响应信息后的第 T 个时隙内的动作和报文传输结果。可理解的，数据 1 子字段用于指示 STA 自上一次成功上报动作信息以来，在第一个时隙内的动作和报文传输结果。数据 T 子字段用于指示 STA 自上一次成功上报动作信息以来，在第 T 个时隙内的动作和报文传输结果。

那么,上述 S103 中,AP 根据 N 个动作信息,确定每个 STA 的第一神经网络的训练结果,可指:AP 根据 N 个动作信息和 N 个报文传输结果信息,确定每个 STA 的第一神经网络的训练结果。可理解的,AP 无需根据动作信息确定报文传输结果信息,可直接使用接收到的动作信息和报文传输信息确定 STA 的第一神经网络的训练结果,从而可减少 AP 处理的复杂度。

可选的,上述 S103 中,AP 根据 N 个动作信息,确定每个 STA 的第一神经网络的训练结果,还可指:AP 根据 N 个动作信息,以及根据 N 个动作信息确定的 N 个报文传输结果信息,确定每个 STA 的第一神经网络的训练结果。也就是说,该实施方式下,即使 STA 上报了报文传输结果信息,AP 仍可采用根据动作信息确定的报文传输结果信息确定第一神经网络的训练结果。

可理解的,当每个 STA 上报动作信息和载波侦听结果信息时,或者上报动作信息和报文传输结果信息时,AP 对 N 个 STA 上报的 N 个动作信息和 N 个载波侦听结果信息,或者对 N 个动作信息和 N 个报文传输结果信息的处理方式和上述信道接入方法 100 中的处理方式相同,不再赘述。例如,当每个 STA 上报动作信息和载波侦听结果信息时,上述 S103 中,STA 的状态信息是根据 STA 的动作信息和载波侦听结果信获得的,第二神经网络的神经网络参数是根据 N 个动作信息和 N 个载波侦听结果信息获得的,奖励函数是根据 N 个动作信息和 N 个载波侦听结果信息确定的。

五. AP 向每个 STA 下发第一神经网络的训练结果的实施方式。

当 N 个 STA 对应的第一神经网络的神经网络参数相同或不相同时,AP 向每个 STA 下发第一神经网络的训练结果的实施方式可不相同。以下阐述几种 AP 向 N 个 STA 下发第一神经网络的可选实施方式:

1.N 个 STA 共享神经网络参数。

可理解的,当 N 个 STA 共享神经网络参数时,AP 将每个 STA 的第一神经网络的训练结果发送给对应的 STA,是指:AP 将第一神经网络的训练结果广播给 N 个 STA。

也就是说,每个 STA 的第一神经网络的神经网络参数相同时,AP 根据 N 个 STA 上报的动作信息确定的每个第一神经网络的训练结果也是相同的,即 AP 根据 N 个 STA 上报的动作信息确定了一个第一神经网络的训练结果。从而,AP 可将确定的第一神经网络的训练结果通过组播的方式下发给 N 个 STA,可减少系统的开销。

2.N 个 STA 中的 S 个 STA 共享神经网络参数。

可理解的,N 个 STA 中的 S 个 STA 共享一个神经网络参数,S 为小于或等于 N 的正整数。那么,AP 将每个 STA 的第一神经网络的训练结果发送给对应的 STA,是指:AP 将 S 个 STA 对应的第一神经网络的训练结果组播给 S 个 STA,以及将 N-S 个第一神经网络的训练结果单播给对应的 STA。

也就是说,当 N 个 STA 中的部分 STA 共享神经网络参数,另外部分 STA 不共享神经网络参数时,AP 将共享神经网络参数的 STA 的第一神经网络的训练结果通过组播下发给该部分的 STA,将不共享神经网络参数的 STA 的第一神经网络的训练结果单播给对应的 STA。该方式也可减少系统的开销。

3.N 个 STA 不共享神经网络参数。

可理解的,N 个 STA 对应的 N 个第一神经网络的神经网络参数均不相同,AP 根据 N 个 STA 上报的信息确定的每个第一神经网络的训练结果也不相同,因此每个第一神经网络的训练结果是单播给对应的 STA 的。

一种可选的实施方式中,每个 STA 可向 AP 上报用于指示自身和其他 STA 是否共享神经

网络参数的信息，这样 AP 可通过 STA 上报的指示信息确定 N 个 STA 中的部分 STA 或全部 STA 是否共享神经网络参数的情况，进而确定向每个 STA 下发第一神经网络的训练结果的实施方式。

一种可选的实施方式中，每个 STA 上报动作信息之前，或者，AP 将每个第一神经网络的训练结果发送给对应的 STA 之前，AP 将每个 STA 对应的第一神经网络的结构下发给每个 STA，使得每个 STA 获得自身第一神经网络的结构。

另一种可选的实施方式中，每个 STA 的第一神经网络是 AP 预定义的，即每个 STA 预先知道自身第一神经网络的结构和第一神经网络的神经网络参数，无需 AP 通过信令下发告知每个 STA，从而可节省 AP 的信令开销。

又一种可选的实施方式中，每个 STA 上报动作信息之前，或者，AP 将每个第一神经网络的训练结果发送给对应的 STA 之前，AP 将多个第一神经网络的结构下发给每个 STA，STA 在确定使用某一个第一神经网络的结构时，将确定的第一神经网络的结构上报给 AP，以使得 AP 获得每个 STA 具体使用的第一神经网络的结构。该方式中，每个 STA 可自行灵活的在 AP 下发的多个第一神经网络的结构中选择自身使用的第一神经网络的结构。

本申请实施例中，每个 STA 可向 AP 请求更新自身第一神经网络的训练结果，那么 AP 可在接收到来自 STA 的请求信息时，将该 STA 的第一神经网络的训练结果发送给该 STA。

针对 N 个 STA 的第一神经网络的训练结果，每个第一神经网络的训练结果携带于第二帧中，第二帧的帧结构可参见图 15 所示。第二帧包括 element ID 子字段、length 子字段、element ID extension 子字段以及训练结果 (neural network parameters or gradients，即神经网络参数或梯度)。该第二帧可以是已有的管理帧，也可以是新增的管理帧，具体实现方式可参见上述第一帧的实施方式，不再赘述。

六、通信装置。

为了实现上述本申请实施例提供的方法中的各功能，AP 或 STA 可以包括硬件结构和/或软件模块，以硬件结构、软件模块、或硬件结构加软件模块的形式来实现上述各功能。上述各功能中的某个功能以硬件结构、软件模块、还是硬件结构加软件模块的方式来执行，取决于技术方案的特定应用和设计约束条件。

如图 16 所示，本申请实施例提供了一种通信装置 1600。该通信装置 1600 可以是 AP 的部件 (例如，集成电路，芯片等等)，也可以是 STA 的部件 (例如，集成电路，芯片等等)。该通信装置 1600 也可以是其他通信单元，用于实现本申请方法实施例中的方法。该通信装置 1600 可以包括：通信单元 1601 和处理单元 1602。可选的，还可以包括存储单元 1603。

在一种可能的设计中，如图 16 中的一个或者多个单元可能由一个或者多个处理器来实现，或者由一个或者多个处理器和存储器来实现；或者由一个或多个处理器和收发器实现；或者由一个或者多个处理器、存储器和收发器实现，本申请实施例对此不作限定。所述处理器、存储器、收发器可以单独设置，也可以集成。

所述通信装置 1600 具备实现本申请实施例描述的 AP 的功能，可选的，通信装置 1600 具备实现本申请实施例描述的 STA 的功能。比如，所述通信装置 1600 包括 AP 执行本申请实施例描述的 AP 涉及步骤所对应的模块或单元或手段 (means)，所述功能或单元或手段 (means) 可以通过软件实现，或者通过硬件实现，也可以通过硬件执行相应的软件实现，还可以通过软件和硬件结合的方式实现。详细可进一步参考前述对应方法实施例中的相应描述。

在一种可能的设计中，一种通信装置 1600 可包括：

通信单元 1601，用于接收 N 个站点 STA 分别上报的动作信息，N 个所述动作信息用于

确定每个 STA 的第一神经网络的训练结果, 所述 N 为正整数;

处理单元 1602, 用于根据 N 个所述动作信息, 确定每个 STA 的第一神经网络的训练结果;

通信单元 1601, 还用于将所述每个 STA 的第一神经网络的训练结果发送给对应的 STA。

一种可选的实施方式中, 所述动作信息是指一段时间的动作, 所述动作为发送或不发送。

一种可选的实现方式中, 处理单元 1602, 还用于: 通过通信单元 1601 接收所述 N 个 STA 分别上报的载波侦听结果信息或报文传输结果信息; 处理单元 1602 根据 N 个所述动作信息, 确定每个 STA 的第一神经网络的训练结果, 具体用于: 根据 N 个所述动作信息和 N 个所述载波侦听结果信息, 确定每个 STA 的第一神经网络的训练结果; 或者, 根据 N 个所述动作信息和 N 个所述报文传输结果信息, 确定每个 STA 的第一神经网络的训练结果。

一种可选的实现方式中, 所述训练结果为神经网络参数或梯度, 所述神经网络参数/梯度用于对应的 STA 更新对应的第一神经网络。

一种可选的实现方式中, 所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中; 所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段, 所述 T 为正整数;

所述时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻, 所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息; 所述数据 1 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的第一个时隙内的动作; 所述数据 T 子字段用于指示 STA 上一次成功接收所述第一响应信息后的第 T 个时隙内的动作。

另一种可选的实现方式中, 所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中; 所述动作详细信息字段包括时间指示子字段、动作 1 子字段和时间 1 子字段、.....、动作 P 子字段和时间 P 子字段, 所述 P 为正整数;

所述时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻, 所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息; 所述动作 1 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的第一个动作, 所述时间 1 子字段用于指示所述动作 1 的持续时间或所述动作 1 的结束时间; 所述动作 P 子字段用于指示 STA 在上一次成功接收所述第一响应信息的时刻距离当前时刻内的第 P 个动作, 所述时间 P 子字段用于指示所述动作 P 的持续时间或所述动作 P 的结束时间。

又一种可选的实现方式中, 所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中; 所述动作详细信息字段包括时间 1 指示子字段和动作 1 子字段、.....、时间 P 指示子字段和动作 P 子字段, 所述 P 为正整数;

所述时间 1 指示子字段用于指示动作 1 的起始时间, 所述动作 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个动作, 所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息; 所述时间 P 指示子字段用于指示动作 P 的起始时间, 所述动作 P 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的时刻距离当前时刻内的第 P 个动作。

又一种可选的实现方式中, 所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中; 所述动作详细信息字段包括时间 1 指示子字段和持续时间 1 子字段、.....、时间 K 指示子字段和持续时间 K 子字段, 所述 K 为正整数;

所述时间 1 指示子字段用于指示动作 1 的开始时间/结束时间, 所述动作 1 是 STA 在上一次成功接收第一响应信息后, 第一次发送报文且未接收到第二响应信息时的发送动作, 所述

第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息，所述第二响应信息是所述 AP 成功接收 STA 发送的报文时发送的响应信息，所述持续时间 1 子字段用于指示所述动作 1 的持续时间；

所述时间 K 指示子字段用于指示动作 K 的开始时间/结束时间，所述动作 K 是 STA 在上一次成功接收所述第一响应信息后，第 K 次发送报文且未接收到所述第二响应信息时的发送动作，所述持续时间 K 子字段用于指示所述动作 K 的持续时间。

又一种可选的实现方式中，所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中；所述动作详细信息字段包括第一时间 1 指示子字段和第二时间 1 指示子字段、……、第一时间 K 指示子字段和第二时间 K 指示子字段，所述 K 为正整数；

所述第一时间 1 指示子字段用于指示动作 1 的起始时间，所述动作 1 是 STA 在上一次成功接收第一响应信息后，第一次发送报文且未接收到第二响应信息时的发送动作，所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息，所述第二响应信息是所述 AP 成功接收 STA 发送的报文时发送的响应信息，所述第二时间 1 指示子字段用于指示所述动作 1 的结束时间；

所述第一时间 K 指示子字段用于指示动作 K 的起始时间，所述动作 K 是 STA 在上一次成功接收所述第一响应信息后，第 K 次发送报文且未接收到所述第二响应信息时的发送动作，所述第二时间 K 指示子字段用于指示所述动作 K 的结束时间。

又一种可选的实现方式中，所述动作信息和所述载波侦听结果信息携带于 STA 上报的第一帧的动作详细信息字段中；所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段，所述 T 为正整数；

所述时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻，所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息；

所述数据 1 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的第一个时隙内的动作和载波侦听结果；所述数据 T 子字段用于指示 STA 上一次成功接收所述第一响应信息后的第 T 个时隙内的动作和载波侦听结果。

又一种可选的实现方式中，所述动作信息和所述报文传输结果携带于 STA 上报的第一帧的动作详细信息字段中；所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段，所述 T 为正整数；

所述时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻，所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息；

所述数据 1 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的第一个时隙内的动作和报文传输结果；所述数据 T 子字段用于指示 STA 上一次成功接收所述第一响应信息后的第 T 个时隙内的动作和报文传输结果。

一种可选的实施方式中，处理单元 1602 根据所述 N 个动作信息，确定每个 STA 的第一神经网络的训练结果，具体用于：将每个 STA 的状态信息输入到对应 STA 的第一神经网络，获得所述第一神经网络的输出值；每个第一神经网络的输出值输入到第二神经网络，获得所述第二神经网络的输出值；所述第二神经网络的输出值用于表征预设时间内的期望奖励；根据所述第二神经网络的输出值和奖励函数训练第三神经网络，通过最小化所述第三神经网络的损失函数，确定所述每个第一神经网络的训练结果；所述第三神经网络包括所述每个第一神经网络和所述第二神经网络；

其中，所述 STA 的状态信息是根据 STA 的动作信息获得的，所述第二神经网络的神经网络

络参数是根据所述 N 个动作信息获得的, 所述奖励函数是根据所述 N 个动作信息确定的; 或者,

所述 STA 的状态信息是根据 STA 的动作信息和载波侦听结果信息获得的, 所述第二神经网络的神经网络参数是根据 N 个所述动作信息和 N 个所述载波侦听结果信息获得的, 所述奖励函数是根据 N 个所述动作信息和 N 个所述载波侦听结果信息确定的; 或者,

所述 STA 的状态信息是根据 STA 的动作信息和报文传输结果信息获得的, 所述第二神经网络的神经网络参数是根据 N 个所述动作信息和 N 个所述报文传输结果信息获得的, 所述奖励函数是根据 N 个所述动作信息和 N 个所述报文传输结果信息确定的。

一种可选的实现方式中, 处理单元 1602 还用于: 在根据所述 N 个动作信息确定第一 STA 发送报文成功时, 将所述奖励函数的值设置为 1, 所述第一 STA 是所述 N 个 STA 中上一次成功接收第二响应信息的时刻距离当前时刻的时间间隔最长的 STA。

另一种可选的实现方式中, 处理单元 1602 还用于: 在根据所述 N 个动作信息确定第二 STA 发送报文成功时, 将所述奖励函数的值设置为第一时长减 1; 所述第二 STA 是所述 N 个 STA 中, 除第一 STA 外的 STA; 所述第一 STA 是所述 N 个 STA 中上一次成功接收第二响应信息的时刻距离当前时刻的时间间隔最长的 STA; 所述第一时长是所述第二 STA 上一次成功接收所述第二响应信息的时刻距离当前时刻的时长。

又一种可选的实现方式中, 处理单元 1602 还用于: 在根据所述 N 个动作信息确定所述 N 个 STA 中的 M 个 STA 在同一时隙上发送报文时, 将所述奖励函数的值设置为-1; 所述 M 为小于或等于所述 N 的正整数。

又一种可选的实现方式中, 处理单元 1602 还用于: 根据所述 N 个动作信息确定所述 N 个 STA 在同一时隙上均未发送报文时, 将所述奖励函数的值设置为 0。

一种可选的实现方式中, 所述 N 个 STA 共享神经网络参数, 所述通信单元 1601 将所述每个 STA 的第一神经网络的训练结果发送给对应的 STA, 具体用于: 将所述第一神经网络的训练结果广播给所述 N 个 STA。

另一种可选的实现方式中, 所述 N 个 STA 中的 S 个 STA 共享神经网络参数, 所述 S 为小于或等于所述 N 的正整数, 所述通信单元 1601 将所述每个 STA 的第一神经网络的训练结果发送给对应的 STA, 具体用于: 将 S 个 STA 对应的第一神经网络的训练结果组播给所述 S 个 STA, 以及将 N-S 个第一神经网络的训练结果单播给对应的 STA。

一种可选的实现方式中, 所述 N 个 STA 不共享神经网络参数时, 所述每个第一神经网络的训练结果是单播给对应的 STA 的。

本申请实施例和上述所示方法实施例基于同一构思, 其带来的技术效果也相同, 具体原理请参照上述所示实施例的描述, 不再赘述。

在另一种可能的设计中, 一种通信装置 1600 可包括:

通信单元 1601, 用于向接入点 AP 上报动作信息, 所述动作信息用于确定处理单元的第一神经网络的训练结果;

通信单元 1601, 还用于接收来自所述 AP 的第一神经网络的训练结果, 所述第一神经网络的训练结果用于更新所述第一神经网络, 以确定所述处理单元是否接入信道;

处理单元 1602, 用于根据所述第一神经网络的训练结果更新所述第一神经网络, 并在侦听到信道空闲时根据更新后的第一神经网络和所述处理单元当前的状态信息确定是否接入信道。

一种可选的实现方式中, 所述动作信息是指一段时间的动作, 所述动作为发送或不发送。

一种可选的实现方式中,通信单元 1601,还用于向所述 AP 上报载波侦听结果信息或报文传输结果信息,所述载波侦听结果信息或所述报文传输结果信息用于确定所述处理单元的第一神经网络的训练结果。

一种可选的实现方式中,所述训练结果为神经网络参数或梯度,所述神经网络参数/梯度用于所述处理单元 1602 更新所述第一神经网络。

一种可选的实现方式中,所述动作信息携带于所述处理单元 1602 上报的第一帧的动作详细信息字段中;所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段,所述 T 为正整数;

所述时间指示子字段用于指示所述处理单元 1602 上一次成功接收第一响应信息的时刻,所述第一响应信息是所述成功接收所述处理单元 1602 发送的动作信息时发送的响应信息;所述数据 1 子字段用于指示所述处理单元 1602 在上一次成功接收所述第一响应信息后的第一个时隙内的动作;所述数据 T 子字段用于指示所述处理单元 1602 上一次成功接收所述第一响应信息后的第 T 个时隙内的动作。

另一种可选的实施方式中,所述动作信息携带于所述处理单元 1602 上报的第一帧的动作详细信息字段中;所述动作详细信息字段包括时间指示子字段、动作 1 子字段和时间 1 子字段、……、动作 P 子字段和时间 P 子字段,所述 P 为正整数;

所述时间指示子字段用于指示所述处理单元 1602 上一次成功接收第一响应信息的时刻,所述第一响应信息是所述 AP 成功接收所述处理单元 1602 发送的动作信息时发送的响应信息;所述动作 1 子字段用于指示所述处理单元 1602 在上一次成功接收所述第一响应信息后的第一个动作,所述时间 1 子字段用于指示所述动作 1 的持续时间或所述动作 1 的结束时间;所述动作 P 子字段用于指示所述处理单元 1602 在上一次成功接收所述第一响应信息的时刻距离当前时刻内的第 P 个动作,所述时间 P 子字段用于指示所述动作 P 的持续时间或所述动作 P 的结束时间。

又一种可选的实现方式中,所述动作信息携带于所述处理单元 1602 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间 1 指示子字段和动作 1 子字段、……、时间 P 指示子字段和动作 P 子字段,所述 P 为正整数;

所述时间 1 指示子字段用于指示动作 1 的起始时间,所述动作 1 子字段用于指示所述处理单元 1602 在上一次成功接收第一响应信息后的第一个动作,所述第一响应信息是所述 AP 成功接收所述 STA 发送的动作信息时发送的响应信息;

所述时间 P 指示子字段用于指示动作 P 的起始时间,所述动作 P 子字段用于指示所述处理单元 1602 在上一次成功接收所述第一响应信息后的时刻距离当前时刻内的第 P 个动作。

又一种可选的实施方式中,所述动作信息携带于所述处理单元 1602 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间 1 指示子字段和持续时间 1 子字段、……、时间 K 指示子字段和持续时间 K 子字段,所述 K 为正整数;

所述时间 1 指示子字段用于指示动作 1 的开始时间/结束时间,所述动作 1 是所述 STA 在上一次成功接收第一响应信息后,第一次发送报文且未接收到第二响应信息时的发送动作,所述第一响应信息是所述 AP 成功接收所述处理单元 1602 发送的动作信息时发送的响应信息,所述第二响应信息是所述 AP 成功接收所述处理单元 1602 发送的报文时发送的响应信息,所述持续时间 1 子字段用于指示所述动作 1 的持续时间;

所述时间 K 指示子字段用于指示动作 K 的开始时间/结束时间, 所述动作 K 是所述处理单元 1602 在上一次成功接收所述第一响应信息后, 第 K 次发送报文且未接收到所述第二响应信息时的发送动作, 所述持续时间 K 子字段用于指示所述动作 K 的持续时间。

又一种可选的实施方式中, 所述动作信息携带于所述处理单元 1602 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括第一时间 1 指示子字段和第二时间 1 指示子字段、.....、第一时间 K 指示子字段和第二时间 K 指示子字段, 所述 K 为正整数;

所述第一时间 1 指示子字段用于指示动作 1 的起始时间, 所述动作 1 是所述处理单元 1602 在上一次成功接收第一响应信息后, 第一次发送报文且未接收到第二响应信息时的发送动作, 所述第一响应信息是所述 AP 成功接收所述处理单元 1602 发送的动作信息时发送的响应信息, 所述第二响应信息是所述 AP 成功接收所述处理单元 1602 发送的报文时发送的响应信息, 所述第二时间 1 指示子字段用于指示所述动作 1 的结束时间;

所述第一时间 K 指示子字段用于指示动作 K 的起始时间, 所述动作 K 是所述处理单元 1602 在上一次成功接收所述第一响应信息后, 第 K 次发送报文且未接收到所述第二响应信息时的发送动作, 所述第二时间 K 指示子字段用于指示所述动作 K 的结束时间。

又一种可选的实施方式中, 所述动作信息和所述载波侦听结果信息携带于所述处理单元 1602 上报的第一帧的动作详细信息字段中; 所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段, 所述 T 为正整数;

所述时间指示子字段用于指示所述处理单元 1602 上一次成功接收第一响应信息的时刻, 所述第一响应信息是所述 AP 成功接收所述处理单元 1602 发送的动作信息时发送的响应信息;

所述数据 1 子字段用于指示所述处理单元 1602 在上一次成功接收所述第一响应信息后的第一个时隙内的动作和载波侦听结果;

所述数据 T 子字段用于指示所述处理单元 1602 上一次成功接收所述第一响应信息后的第 T 个时隙内的动作和载波侦听结果。

又一种可选的实施方式中, 所述动作信息和所述报文传输结果携带于所述处理单元 1602 上报的第一帧的动作详细信息字段中; 所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段, 所述 T 为正整数;

所述时间指示子字段用于指示处理单元 1602 上一次成功接收第一响应信息的时刻, 所述第一响应信息是所述 AP 成功接收所述处理单元 1602 发送的动作信息时发送的响应信息;

所述数据 1 子字段用于指示所述处理单元 1602 在上一次成功接收所述第一响应信息后的第一个时隙内的动作和报文传输结果;

所述数据 T 子字段用于指示所述处理单元 1602 上一次成功接收所述第一响应信息后的第 T 个时隙内的动作和报文传输结果。

一种可选的实施方式中, 所述处理单元 1602 根据所述第一神经网络的训练结果更新所述第一神经网络, 并在侦听到信道空闲时根据更新后的所述第一神经网络和所述处理单元当前的状态信息确定是否接入信道, 具体用于: 将所述处理单元当前的状态信息输入到更新后的所述第一神经网络, 输出第一值和第二值; 所述第一值用于表征接入信道获得的预期奖励, 所述第二值用于表征不接入信道获得的预期奖励; 在所述第一值大于所述第二值时, 确定接入信道; 在所述第一值小于所述第二值时, 确定不接入信道。

本申请实施例和上述所示方法实施例基于同一构思, 其带来的技术效果也相同, 具体原

理请参照上述所示实施例的描述，不再赘述。

本申请实施例还提供一种通信装置 1700，图 17 为通信装置 1700 的结构示意图。所述通信装置 1700 可以是 AP 或 STA，也可以是支持 AP 实现上述方法的芯片、芯片系统、或处理器等，还可以是支持 STA 实现上述方法的芯片、芯片系统、或处理器等。该装置可用于实现上述方法实施例中描述的方法，具体可以参见上述方法实施例中的说明。

所述通信装置 1700 可以包括一个或多个处理器 1701。所述处理器 1701 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、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、分立门或者晶体管逻辑器件、分立硬件组件或中央处理器（Central Processing Unit, CPU）。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，基站、基带芯片，终端、终端芯片，DU 或 CU 等）进行控制，执行软件程序，处理软件程序的数据。

可选的，所述通信装置 1700 中可以包括一个或多个存储器 1702，其上可以存有指令 1704，所述指令可在所述处理器 1701 上被运行，使得所述通信装置 1700 执行上述方法实施例中描述的方法。可选的，所述存储器 1702 中还可以存储有数据。所述处理器 1701 和存储器 1702 可以单独设置，也可以集成在一起。

存储器 1702 可包括但不限于硬盘（hard disk drive, HDD）或固态硬盘（solid-state drive, SSD）等非易失性存储器，随机存储记忆体（Random Access Memory, RAM）、可擦除可编程只读存储器（Erasable Programmable ROM, EPROM）、只读存储器（Read-Only Memory, ROM）或便携式只读存储器（Compact Disc Read-Only Memory, CD-ROM）等等。

可选的，所述通信装置 1700 还可以包括收发器 1705、天线 1706。所述收发器 1705 可以称为收发单元、收发机、或收发电路等，用于实现收发功能。收发器 1705 可以包括接收器和发送器，接收器可以称为接收机或接收电路等，用于实现接收功能；发送器可以称为发送机或发送电路等，用于实现发送功能。

所述通信装置 1700 为 AP：收发器 1705 用于执行上述信道接入方法 100 中的 S102、S104，处理器 1701 用于执行信道接入方法 100 中的 S103。

所述通信装置 1700 为 STA：处理器 1701 用于执行信道接入方法 100 中的 S106。收发器 1705 用于执行信道接入方法 100 中的 S101、S105。

另一种可能的设计中，处理器 1701 中可以包括用于实现接收和发送功能的收发器。例如该收发器可以是收发电路，或者是接口，或者是接口电路。用于实现接收和发送功能的收发电路、接口或接口电路可以是分开的，也可以集成在一起。上述收发电路、接口或接口电路可以用于代码/数据的读写，或者，上述收发电路、接口或接口电路可以用于信号的传输或传递。

又一种可能的设计中，可选的，处理器 1701 可以存有指令 1703，指令 1703 在处理器 1701 上运行，可使得所述通信装置 1700 执行上述方法实施例中描述的方法。指令 1703 可能固化在处理器 1701 中，该种情况下，处理器 1701 可能由硬件实现。

又一种可能的设计中，通信装置 1700 可以包括电路，所述电路可以实现前述方法实施例中发送或接收或者通信的功能。本申请实施例中描述的处理器和收发器可实现在集成电路（integrated circuit, IC）、模拟 IC、射频集成电路 RFIC、混合信号 IC、专用集成电路（application specific integrated circuit, ASIC）、印刷电路板（printed circuit board, PCB）、电子设备等上。该处理器和收发器也可以用各种 IC 工艺技术来制造，例如互补金属氧化物半导体（complementary metal oxide semiconductor, CMOS）、N 型金属氧化物半导体

(nMetal-oxide-semiconductor, NMOS)、P 型金属氧化物半导体 (positive channel metal oxide semiconductor, PMOS)、双极结型晶体管 (Bipolar Junction Transistor, BJT)、双极 CMOS (BiCMOS)、硅锗 (SiGe)、砷化镓 (GaAs) 等。

本申请实施例和上述信道接入方法 100 所示方法实施例基于同一构思, 其带来的技术效果也相同, 具体原理请参照上述信道接入方法 100 所示实施例的描述, 不再赘述。

本申请还提供了一种计算机可读存储介质, 用于储存计算机软件指令, 当所述指令被通信装置执行时, 实现上述任一方法实施例的功能。

本申请还提供了一种计算机程序产品, 用于储存计算机软件指令, 当所述指令被通信装置执行时, 实现上述任一方法实施例的功能。

本申请还提供了一种计算机程序, 当其在计算机上运行时, 实现上述任一方法实施例的功能。

上述实施例中, 可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时, 可以全部或部分地以计算机程序产品的形式实现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机指令时, 全部或部分地产生按照本申请实施例所述的交互或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中, 或者从一个计算机可读存储介质向另一个计算机可读存储介质传输, 例如, 所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线(例如同轴电缆、光纤、数字用户线(digital subscriber line, DSL))或无线(例如红外、无线、微波等)方式向另一个网站站点、计算机、服务器或数据中心进行传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。所述可用介质可以是磁性介质(例如, 软盘、硬盘、磁带)、光介质(例如, 高密度数字视频光盘(digital video disc, DVD))、或者半导体介质(例如, 固态硬盘(solid state drive, SSD))等。

以上所述, 仅为本申请的具体实施方式, 但本申请的保护范围并不局限于此, 任何熟悉本技术领域的技术人员在本申请揭露的技术范围内, 可轻易想到变化或替换, 都应涵盖在本申请的保护范围之内。因此, 本申请的保护范围应以所述权利要求的保护范围为准。

权 利 要 求 书

1.一种信道接入方法，其特征在于，所述方法包括：

接入点 AP 接收 N 个站点 STA 分别上报的动作信息，N 个所述动作信息用于确定每个 STA 的第一神经网络的训练结果，所述 N 为正整数；

所述 AP 根据 N 个所述动作信息，确定每个 STA 的第一神经网络的训练结果；

所述 AP 将所述每个 STA 的第一神经网络的训练结果发送给对应的 STA。

2.根据权利要求 1 所述的方法，其特征在于，所述动作信息是指一段时间的动作，所述动作为发送或不发送。

3.根据权利要求 1 或 2 所述的方法，其特征在于，所述方法还包括：

所述 AP 接收所述 N 个 STA 分别上报的载波侦听结果信息或报文传输结果信息；

所述 AP 根据 N 个所述动作信息，确定每个 STA 的第一神经网络的训练结果，包括：

所述 AP 根据 N 个所述动作信息和 N 个所述载波侦听结果信息，确定每个 STA 的第一神经网络的训练结果；或者，

所述 AP 根据 N 个所述动作信息和 N 个所述报文传输结果信息，确定每个 STA 的第一神经网络的训练结果。

4.根据权利要求 1 至 3 任一项所述的方法，其特征在于，所述训练结果为神经网络参数或梯度，所述神经网络参数/梯度用于 STA 更新第一神经网络。

5.根据权利要求 1 或 2 所述的方法，其特征在于，所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中；

所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段，所述 T 为正整数；

所述时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻，所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息；

所述数据 1 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的第一个时隙内的动作；

所述数据 T 子字段用于指示 STA 上一次成功接收所述第一响应信息后的第 T 个时隙内的动作。

6.根据权利要求 1 或 2 所述的方法，其特征在于，所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中；

所述动作详细信息字段包括时间指示子字段、动作 1 子字段和时间 1 子字段、……、动作 P 子字段和时间 P 子字段，所述 P 为正整数；

所述时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻，所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息；

所述动作 1 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的第一个动作，

所述时间 1 子字段用于指示所述动作 1 的持续时间或所述动作 1 的结束时间；

所述动作 P 子字段用于指示 STA 在上一次成功接收所述第一响应信息的时刻距离当前时刻内的第 P 个动作，所述时间 P 子字段用于指示所述动作 P 的持续时间或所述动作 P 的结束时间。

7. 根据权利要求 1 或 2 所述的方法，其特征在于，所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中；

所述动作详细信息字段包括时间 1 指示子字段和动作 1 子字段、……、时间 P 指示子字段和动作 P 子字段，所述 P 为正整数；

所述时间 1 指示子字段用于指示动作 1 的起始时间，所述动作 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个动作，所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息；

所述时间 P 指示子字段用于指示动作 P 的起始时间，所述动作 P 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的时刻距离当前时刻内的第 P 个动作。

8. 根据权利要求 1 或 2 所述的方法，其特征在于，所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中；

所述动作详细信息字段包括时间 1 指示子字段和持续时间 1 子字段、……、时间 K 指示子字段和持续时间 K 子字段，所述 K 为正整数；

所述时间 1 指示子字段用于指示动作 1 的开始时间/结束时间，所述动作 1 是 STA 在上一次成功接收第一响应信息后，第一次发送报文且未接收到第二响应信息时的发送动作，所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息，所述第二响应信息是所述 AP 成功接收 STA 发送的报文时发送的响应信息，所述持续时间 1 子字段用于指示所述动作 1 的持续时间；

所述时间 K 指示子字段用于指示动作 K 的开始时间/结束时间，所述动作 K 是 STA 在上一次成功接收所述第一响应信息后，第 K 次发送报文且未接收到所述第二响应信息时的发送动作，所述持续时间 K 子字段用于指示所述动作 K 的持续时间。

9. 根据权利要求 1 或 2 所述的方法，其特征在于，所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中；

所述动作详细信息字段包括第一时间 1 指示子字段和第二时间 1 指示子字段、……、第一时间 K 指示子字段和第二时间 K 指示子字段，所述 K 为正整数；

所述第一时间 1 指示子字段用于指示动作 1 的起始时间，所述动作 1 是 STA 在上一次成功接收第一响应信息后，第一次发送报文且未接收到第二响应信息时的发送动作，所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息，所述第二响应信息是所述 AP 成功接收 STA 发送的报文时发送的响应信息，所述第二时间 1 指示子字段用于指示所述动作 1 的结束时间；

所述第一时间 K 指示子字段用于指示动作 K 的起始时间，所述动作 K 是 STA 在上一次成功接收所述第一响应信息后，第 K 次发送报文且未接收到所述第二响应信息时的发送动作，所述第二时间 K 指示子字段用于指示所述动作 K 的结束时间。

10.根据权利要求3所述的方法,其特征在于,所述动作信息和所述载波侦听结果信息携带于STA上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间指示子字段、数据1子字段至数据T子字段,所述T为正整数;

所述时间指示子字段用于指示STA上一次成功接收第一响应信息的时刻,所述第一响应信息是所述AP成功接收STA发送的动作信息时发送的响应信息;

所述数据1子字段用于指示STA在上一次成功接收所述第一响应信息后的第一个时隙内的动作和载波侦听结果;

所述数据T子字段用于指示STA在上一次成功接收所述第一响应信息后的第T个时隙内的动作和载波侦听结果。

11.根据权利要求3所述的方法,其特征在于,所述动作信息和所述报文传输结果信息携带于STA上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间指示子字段、数据1子字段至数据T子字段,所述T为正整数;

所述时间指示子字段用于指示STA上一次成功接收第一响应信息的时刻,所述第一响应信息是所述AP成功接收STA发送的动作信息时发送的响应信息;

所述数据1子字段用于指示STA在上一次成功接收所述第一响应信息后的第一个时隙内的动作和报文传输结果;

所述数据T子字段用于指示STA在上一次成功接收所述第一响应信息后的第T个时隙内的动作和报文传输结果。

12.根据权利要求1至11任一项所述的方法,其特征在于,所述AP根据N个所述动作信息,确定每个STA的第一神经网络的训练结果,包括:

所述AP将每个STA的状态信息输入到对应STA的第一神经网络,获得所述第一神经网络的输出值;

所述AP将每个第一神经网络的输出值输入到第二神经网络,获得所述第二神经网络的输出值,所述第二神经网络的输出值用于表征预设时间内的期望奖励;

所述AP根据所述第二神经网络的输出值和奖励函数训练第三神经网络,通过最小化所述第三神经网络的损失函数,确定所述每个第一神经网络的训练结果,所述第三神经网络包括所述每个第一神经网络和所述第二神经网络;

其中,所述STA的状态信息是根据STA的动作信息获得的,所述第二神经网络的神经网络参数是根据N个所述动作信息获得的,所述奖励函数是根据N个所述动作信息确定的;或者,

所述STA的状态信息是根据STA的动作信息和载波侦听结果信息获得的,所述第二神经网络的神经网络参数是根据N个所述动作信息和N个所述载波侦听结果信息获得的,所述奖励函数是根据N个所述动作信息和N个所述载波侦听结果信息确定的;或者,

所述STA的状态信息是根据STA的动作信息和报文传输结果信息获得的,所述第二神经网络的神经网络参数是根据N个所述动作信息和N个所述报文传输结果信息获得的,所述奖励函数是根据N个所述动作信息和N个所述报文传输结果信息确定的。

13.根据权利要求 12 所述的方法，其特征在于，所述方法还包括：

所述 AP 在根据 N 个所述动作信息确定第一 STA 发送报文成功时，将所述奖励函数的值设置为 1，所述第一 STA 是所述 N 个 STA 中上一次成功接收第二响应信息的时刻距离当前时刻的时间间隔最长的 STA。

14.根据权利要求 12 所述的方法，其特征在于，所述方法还包括：

所述 AP 在根据 N 个所述动作信息确定第二 STA 发送报文成功时，将所述奖励函数的值设置为第一时长减 1；

所述第二 STA 是所述 N 个 STA 中，除第一 STA 外的 STA；所述第一 STA 是所述 N 个 STA 中上一次成功接收第二响应信息的时刻距离当前时刻的时间间隔最长的 STA；

所述第一时长是所述第二 STA 上一次成功接收所述第二响应信息的时刻距离当前时刻的时长。

15.根据权利要求 12 所述的方法，其特征在于，所述方法还包括：

所述 AP 在根据 N 个所述动作信息确定所述 N 个 STA 中的 M 个 STA 在同一时隙上发送报文时，将所述奖励函数的值设置为-1，所述 M 为小于或等于所述 N 的正整数。

16.根据权利要求 12 所述的方法，其特征在于，所述方法还包括：

所述 AP 在根据 N 个所述动作信息确定所述 N 个 STA 在同一时隙上均未发送报文时，将所述奖励函数的值设置为 0。

17.根据权利要求 1 至 16 任一项所述的方法，其特征在于，所述 N 个 STA 共享神经网络参数；所述 AP 将所述每个 STA 的第一神经网络的训练结果发送给对应的 STA，包括：

所述 AP 将所述第一神经网络的训练结果广播给所述 N 个 STA。

18.根据权利要求 1 至 16 任一项所述的方法，其特征在于，所述 N 个 STA 中的 S 个 STA 共享神经网络参数，所述 S 为小于或等于所述 N 的正整数；所述 AP 将所述每个 STA 的第一神经网络的训练结果发送给对应的 STA，包括：

所述 AP 将 S 个 STA 对应的第一神经网络的训练结果组播给所述 S 个 STA，以及将 N-S 个第一神经网络的训练结果单播给对应的 STA。

19.根据权利要求 1 至 16 任一项所述的方法，其特征在于，

所述 N 个 STA 不共享神经网络参数时，所述每个第一神经网络的训练结果是单播给对应的 STA 的。

20.一种信道接入方法，其特征在于，所述方法包括：

站点 STA 向接入点 AP 上报动作信息，所述动作信息用于确定所述 STA 的第一神经网络的训练结果；

所述 STA 接收来自所述 AP 的第一神经网络的训练结果，所述第一神经网络的训练结果用于更新所述第一神经网络，以确定所述 STA 是否接入信道；

所述 STA 根据所述第一神经网络的训练结果更新所述第一神经网络，并在侦听到信道空

闲时根据更新后的第一神经网络和所述 STA 当前的状态信息确定是否接入信道。

21.根据权利要求 20 所述的方法,其特征在于,所述动作信息是指一段时间的动作,所述动作为发送或不发送。

22.根据权利要求 20 或 21 所述的方法,其特征在于,所述方法还包括:

所述 STA 向所述 AP 上报载波侦听结果信息或报文传输结果信息,所述载波侦听结果信息或所述报文传输结果信息用于确定所述 STA 的第一神经网络的训练结果。

23.根据权利要求 20 至 22 任一项所述的方法,其特征在于,所述训练结果为神经网络参数或梯度;

所述神经网络参数/梯度用于所述 STA 更新所述第一神经网络。

24.根据权利要求 20 或 21 所述的方法,其特征在于,所述动作信息携带于所述 STA 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段,所述 T 为正整数;

所述时间指示子字段用于指示所述 STA 上一次成功接收第一响应信息的时刻,所述第一响应信息是所述成功接收所述 STA 发送的动作信息时发送的响应信息;

所述数据 1 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息后的第一个时隙内的动作;

所述数据 T 子字段用于指示所述 STA 上一次成功接收所述第一响应信息后的第 T 个时隙内的动作。

25.根据权利要求 20 或 21 所述的方法,其特征在于,所述动作信息携带于所述 STA 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间指示子字段、动作 1 子字段和时间 1 子字段、.....、动作 P 子字段和时间 P 子字段,所述 P 为正整数;

所述时间指示子字段用于指示所述 STA 上一次成功接收第一响应信息的时刻,所述第一响应信息是所述 AP 成功接收所述 STA 发送的动作信息时发送的响应信息;

所述动作 1 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息后的第一个动作,所述持续时间 1 子字段用于指示所述动作 1 的持续时间或所述动作 1 的结束时间;

所述动作 P 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息的时刻距离当前时刻内的第 P 个动作,所述持续时间 P 子字段用于指示所述动作 P 的持续时间或所述动作 P 的结束时间。

26.根据权利要求 20 或 21 所述的方法,其特征在于,所述动作信息携带于所述 STA 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间 1 指示子字段和动作 1 子字段、.....、时间 P 指示子字段和动作 P 子字段,所述 P 为正整数;

所述时间 1 指示子字段用于指示动作 1 的起始时间,所述动作 1 子字段用于指示所述 STA

在上一次成功接收第一响应信息后的第一个动作,所述第一响应信息是所述 AP 成功接收所述 STA 发送的动作信息时发送的响应信息;

所述时间 P 指示子字段用于指示动作 P 的起始时间,所述动作 P 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息后的时刻距离当前时刻内的第 P 个动作。

27.根据权利要求 20 或 21 所述的方法,其特征在于,所述动作信息携带于所述 STA 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间 1 指示子字段和持续时间 1 子字段、.....、时间 K 指示子字段和持续时间 K 子字段,所述 K 为正整数;

所述时间 1 指示子字段用于指示动作 1 的开始时间/结束时间,所述动作 1 是所述 STA 在上一次成功接收第一响应信息后,第一次发送报文且未接收到第二响应信息时的发送动作,所述第一响应信息是所述 AP 成功接收所述 STA 发送的动作信息时发送的响应信息,所述第二响应信息是所述 AP 成功接收所述 STA 发送的报文时发送的响应信息,所述持续时间 1 子字段用于指示所述动作 1 的持续时间;

所述时间 K 指示子字段用于指示动作 K 的开始时间/结束时间,所述动作 K 是所述 STA 在上一次成功接收所述第一响应信息后,第 K 次发送报文且未接收到所述第二响应信息时的发送动作,所述持续时间 K 子字段用于指示所述动作 K 的持续时间。

28.根据权利要求 20 或 21 所述的方法,其特征在于,所述动作信息携带于所述 STA 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括第一时间 1 指示子字段和第一时间 K 指示子字段、.....、第一时间 K 指示子字段和第一时间 K 指示子字段,所述 K 为正整数;

所述第一时间 1 指示子字段用于指示动作 1 的起始时间,所述动作 1 是所述 STA 在上一次成功接收第一响应信息后,第一次发送报文且未接收到第二响应信息时的发送动作,所述第一响应信息是所述 AP 成功接收所述 STA 发送的动作信息时发送的响应信息,所述第二响应信息是所述 AP 成功接收所述 STA 发送的报文时发送的响应信息,所述第一时间 1 指示子字段用于指示所述动作 1 的结束时间;

所述第一时间 K 指示子字段用于指示动作 K 的起始时间,所述动作 K 是所述 STA 在上一次成功接收所述第一响应信息后,第 K 次发送报文且未接收到所述第二响应信息时的发送动作,所述第一时间 K 指示子字段用于指示所述动作 K 的结束时间。

29.根据权利要求 22 所述的方法,其特征在于,所述动作信息和所述载波侦听结果信息携带于所述 STA 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段,所述 T 为正整数;

所述时间指示子字段用于指示所述 STA 上一次成功接收第一响应信息的时刻,所述第一响应信息是所述 AP 成功接收所述 STA 发送的动作信息时发送的响应信息;

所述数据 1 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息后的第一个时隙内的动作和载波侦听结果;

所述数据 T 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息后的第 T 个时隙内的动作和载波侦听结果。

30.根据权利要求 22 所述的方法,其特征在于,所述动作信息和所述报文传输结果信息携带于所述 STA 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段,所述 T 为正整数;

所述时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻,所述第一响应信息是所述 AP 成功接收所述 STA 发送的动作信息时发送的响应信息;

所述数据 1 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息后的第一个时隙内的动作和报文传输结果;

所述数据 T 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息后的第 T 个时隙内的动作和报文传输结果。

31.根据权利要求 20 至 30 任一项所述的方法,其特征在于,所述 STA 根据所述第一神经网络的训练结果更新所述第一神经网络,并在侦听到信道空闲时根据更新后的所述第一神经网络和所述 STA 当前的状态信息确定是否接入信道,包括:

所述 STA 将该 STA 当前的状态信息输入到更新后的所述第一神经网络,输出第一值和第二值,所述第一值用于表征接入信道获得的预期奖励,所述第二值用于表征不接入信道获得的预期奖励;

所述 STA 在所述第一值大于所述第二值时,确定接入信道;

所述 STA 在所述第一值小于所述第二值时,确定不接入信道。

32.一种通信装置,其特征在于,所述通信装置包括:

通信单元,用于接收 N 个站点 STA 分别上报的动作信息,N 个所述动作信息用于确定每个 STA 的第一神经网络的训练结果,所述 N 为正整数;

处理单元,用于根据 N 个所述动作信息,确定每个 STA 的第一神经网络的训练结果;

所述通信单元,还用于将所述每个 STA 的第一神经网络的训练结果发送给对应的 STA。

33.根据权利要求 32 所述的装置,其特征在于,所述动作信息是指一段时间的动作,所述动作为发送或不发送。

34.根据权利要求 32 或 33 所述的装置,其特征在于,

所述通信单元,还用于接收所述 N 个 STA 分别上报的载波侦听结果信息或报文传输结果信息;

所述处理单元根据 N 个所述动作信息,确定每个 STA 的第一神经网络的训练结果,具体用于:

根据 N 个所述动作信息和 N 个所述载波侦听结果信息,确定每个 STA 的第一神经网络的训练结果;或者,

根据 N 个所述动作信息和 N 个所述报文传输结果信息,确定每个 STA 的第一神经网络的训练结果。

35.根据权利要求 32 至 34 任一项所述的装置,其特征在于,所述训练结果为神经网络参

数或梯度, 所述神经网络参数/梯度用于 STA 更新第一神经网络。

36. 根据权利要求 32 或 33 所述的装置, 其特征在于, 所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段, 所述 T 为正整数;

所述时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻, 所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息;

所述数据 1 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的第一个时隙内的动作;

所述数据 T 子字段用于指示 STA 上一次成功接收所述第一响应信息后的第 T 个时隙内的动作。

37. 根据权利要求 32 或 33 所述的装置, 其特征在于, 所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间指示子字段、动作 1 子字段和时间 1 子字段、.....、动作 P 子字段和时间 P 子字段, 所述 P 为正整数;

所述时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻, 所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息;

所述动作 1 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的第一个动作, 所述时间 1 子字段用于指示所述动作 1 的持续时间或所述动作 1 的结束时间;

所述动作 P 子字段用于指示 STA 在上一次成功接收所述第一响应信息的时刻距离当前时刻内的第 P 个动作, 所述时间 P 子字段用于指示所述动作 P 的持续时间或所述动作 P 的结束时间。

38. 根据权利要求 32 或 33 所述的装置, 其特征在于, 所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间 1 指示子字段和动作 1 子字段、.....、时间 P 指示子字段和动作 P 子字段, 所述 P 为正整数;

所述时间 1 指示子字段用于指示动作 1 的起始时间, 所述动作 1 子字段用于指示 STA 在上一次成功接收第一响应信息后的第一个动作, 所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息;

所述时间 P 指示子字段用于指示动作 P 的起始时间, 所述动作 P 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的时刻距离当前时刻内的第 P 个动作。

39. 根据权利要求 32 或 33 所述的装置, 其特征在于, 所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间 1 指示子字段和持续时间 1 子字段、.....、时间 K 指示子字段和持续时间 K 子字段, 所述 K 为正整数;

所述时间 1 指示子字段用于指示动作 1 的开始时间/结束时间, 所述动作 1 是 STA 在上一次成功接收第一响应信息后, 第一次发送报文且未接收到第二响应信息时的发送动作, 所

述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息，所述第二响应信息是所述 AP 成功接收 STA 发送的报文时发送的响应信息，所述持续时间 1 子字段用于指示所述动作 1 的持续时间；

所述时间 K 指示子字段用于指示动作 K 的开始时间/结束时间，所述动作 K 是 STA 在上一次成功接收所述第一响应信息后，第 K 次发送报文且未接收到所述第二响应信息时的发送动作，所述持续时间 K 子字段用于指示所述动作 K 的持续时间。

40. 根据权利要求 32 或 33 所述的装置，其特征在于，所述动作信息携带于 STA 上报的第一帧的动作详细信息字段中；

所述动作详细信息字段包括第一时间 1 指示子字段和第一时间 1 指示子字段、……、第一时间 K 指示子字段和第一时间 K 指示子字段，所述 K 为正整数；

所述第一时间 1 指示子字段用于指示动作 1 的起始时间，所述动作 1 是 STA 在上一次成功接收第一响应信息后，第一次发送报文且未接收到第二响应信息时的发送动作，所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息，所述第二响应信息是所述 AP 成功接收 STA 发送的报文时发送的响应信息，所述第一时间 1 指示子字段用于指示所述动作 1 的结束时间；

所述第一时间 K 指示子字段用于指示动作 K 的起始时间，所述动作 K 是 STA 在上一次成功接收所述第一响应信息后，第 K 次发送报文且未接收到所述第二响应信息时的发送动作，所述第一时间 K 指示子字段用于指示所述动作 K 的结束时间。

41. 根据权利要求 34 所述的装置，其特征在于，所述动作信息和所述载波侦听结果信息携带于 STA 上报的第一帧的动作详细信息字段中；

所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段，所述 T 为正整数；

所述时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻，所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息；

所述数据 1 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的第一个时隙内的动作和载波侦听结果；

所述数据 T 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的第 T 个时隙内的动作和载波侦听结果。

42. 根据权利要求 34 所述的装置，其特征在于，所述动作信息和所述报文传输结果信息携带于 STA 上报的第一帧的动作详细信息字段中；

所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段，所述 T 为正整数；

所述时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻，所述第一响应信息是所述 AP 成功接收 STA 发送的动作信息时发送的响应信息；

所述数据 1 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的第一个时隙内的动作和报文传输结果；

所述数据 T 子字段用于指示 STA 在上一次成功接收所述第一响应信息后的第 T 个时隙内的动作和报文传输结果。

43.根据权利要求 32 至 42 任一项所述的装置,其特征在于,所述处理单元根据 N 个所述动作信息,确定每个 STA 的第一神经网络的训练结果,具体用于:

将每个 STA 的状态信息输入到对应 STA 的第一神经网络,获得所述第一神经网络的输出值;

将每个第一神经网络的输出值输入到第二神经网络,获得所述第二神经网络的输出值,所述第二神经网络的输出值用于表征预设时间内的期望奖励;

根据所述第二神经网络的输出值和奖励函数训练第三神经网络,通过最小化所述第三神经网络的损失函数,确定所述每个第一神经网络的训练结果,所述第三神经网络包括所述每个第一神经网络和所述第二神经网络;

其中,所述 STA 的状态信息是根据 STA 的动作信息获得的,所述第二神经网络的神经网络参数是根据 N 个所述动作信息获得的,所述奖励函数是根据 N 个所述动作信息确定的;或者,

所述 STA 的状态信息是根据 STA 的动作信息和载波侦听结果信息获得的,所述第二神经网络的神经网络参数是根据 N 个所述动作信息和 N 个所述载波侦听结果信息获得的,所述奖励函数是根据 N 个所述动作信息和 N 个所述载波侦听结果信息确定的;或者,

所述 STA 的状态信息是根据 STA 的动作信息和报文传输结果信息获得的,所述第二神经网络的神经网络参数是根据 N 个所述动作信息和 N 个所述报文传输结果信息获得的,所述奖励函数是根据 N 个所述动作信息和 N 个所述报文传输结果信息确定的。

44.根据权利要求 43 所述的装置,其特征在于,所述处理单元还用于:

在根据 N 个所述动作信息确定第一 STA 发送报文成功时,将所述奖励函数的值设置为 1,所述第一 STA 是所述 N 个 STA 中上一次成功接收第二响应信息的时刻距离当前时刻的时间间隔最长的 STA。

45.根据权利要求 43 所述的装置,其特征在于,所述处理单元还用于:

在根据 N 个所述动作信息确定第二 STA 发送报文成功时,将所述奖励函数的值设置为第一时长减 1;

所述第二 STA 是所述 N 个 STA 中,除第一 STA 外的 STA;所述第一 STA 是所述 N 个 STA 中上一次成功接收第二响应信息的时刻距离当前时刻的时间间隔最长的 STA;

所述第一时长是所述第二 STA 上一次成功接收所述第二响应信息的时刻距离当前时刻的时长。

46.根据权利要求 43 所述的装置,其特征在于,所述处理单元还用于:

在根据 N 个所述动作信息确定所述 N 个 STA 中的 M 个 STA 在同一时隙上发送报文时,将所述奖励函数的值设置为-1,所述 M 为小于或等于所述 N 的正整数。

47.根据权利要求 43 所述的装置,其特征在于,所述处理单元还用于:

在根据 N 个所述动作信息确定所述 N 个 STA 在同一时隙上均未发送报文时,将所述奖励函数的值设置为 0。

48.根据权利要求 32 至 47 任一项所述的装置,其特征在于,所述 N 个 STA 共享神经网络参数;所述处理单元将所述每个 STA 的第一神经网络的训练结果发送给对应的 STA,具体用于:

将所述第一神经网络的训练结果广播给所述 N 个 STA。

49.根据权利要求 32 至 47 任一项所述的装置,其特征在于,所述 N 个 STA 中的 S 个 STA 共享神经网络参数,所述 S 为小于或等于所述 N 的正整数;所述处理单元 P 将所述每个 STA 的第一神经网络的训练结果发送给对应的 STA,具体用于:

将 S 个 STA 对应的第一神经网络的训练结果组播给所述 S 个 STA,以及将 N-S 个第一神经网络的训练结果单播给对应的 STA。

50.根据权利要求 32 至 47 任一项所述的方法,其特征在于,

所述 N 个 STA 不共享神经网络参数时,所述每个第一神经网络的训练结果是单播给对应的 STA 的。

51.一种通信装置,其特征在于,所述通信装置包括:

通信单元,用于向接入点 AP 上报动作信息,所述动作信息用于确定处理单元的第一神经网络的训练结果;

通信单元,还用于接收来自所述 AP 的第一神经网络的训练结果,所述第一神经网络的训练结果用于更新所述第一神经网络,以确定所述处理单元是否接入信道;

所述处理单元,用于根据所述第一神经网络的训练结果更新所述第一神经网络,并在侦听到信道空闲时根据更新后的第一神经网络和所述处理单元当前的状态信息确定是否接入信道。

52.根据权利要求 51 所述的装置,其特征在于,所述动作信息是指一段时间的动作,所述动作为发送或不发送。

53.根据权利要求 51 或 52 所述的装置,其特征在于,所述通信单元还用于:

向所述 AP 上报载波侦听结果信息或报文传输结果信息,所述载波侦听结果信息或所述报文传输结果信息用于确定所述 STA 的第一神经网络的训练结果。

54.根据权利要求 51 至 53 任一项所述的装置,其特征在于,所述训练结果为神经网络参数或梯度;

所述神经网络参数/梯度用于所述 STA 更新所述第一神经网络。

55.根据权利要求 51 或 52 所述的装置,其特征在于,所述动作信息携带于所述 STA 上报的第一帧的动作详细信息字段中;

所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段,所述 T 为正整数;

所述时间指示子字段用于指示所述 STA 上一次成功接收第一响应信息的时刻,所述第一响应信息是所述成功接收所述 STA 发送的动作信息时发送的响应信息;

所述数据 1 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息后的第一个时隙内的动作；

所述数据 T 子字段用于指示所述 STA 上一次成功接收所述第一响应信息后的第 T 个时隙内的动作。

56. 根据权利要求 51 或 52 所述的装置，其特征在于，所述动作信息携带于所述 STA 上报的第一帧的动作详细信息字段中；

所述动作详细信息字段包括时间指示子字段、动作 1 子字段和时间 1 子字段、……、动作 P 子字段和时间 P 子字段，所述 P 为正整数；

所述时间指示子字段用于指示所述 STA 上一次成功接收第一响应信息的时刻，所述第一响应信息是所述 AP 成功接收所述 STA 发送的动作信息时发送的响应信息；

所述动作 1 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息后的第一个动作，所述持续时间 1 子字段用于指示所述动作 1 的持续时间或所述动作 1 的结束时间；

所述动作 P 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息的时刻距离当前时刻内的第 P 个动作，所述持续时间 P 子字段用于指示所述动作 P 的持续时间或所述动作 P 的结束时间。

57. 根据权利要求 51 或 52 所述的装置，其特征在于，所述动作信息携带于所述 STA 上报的第一帧的动作详细信息字段中；

所述动作详细信息字段包括时间 1 指示子字段和动作 1 子字段、……、时间 P 指示子字段和动作 P 子字段，所述 P 为正整数；

所述时间 1 指示子字段用于指示动作 1 的起始时间，所述动作 1 子字段用于指示所述 STA 在上一次成功接收第一响应信息后的第一个动作，所述第一响应信息是所述 AP 成功接收所述 STA 发送的动作信息时发送的响应信息；

所述时间 P 指示子字段用于指示动作 P 的起始时间，所述动作 P 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息后的时刻距离当前时刻内的第 P 个动作。

58. 根据权利要求 51 或 52 所述的装置，其特征在于，所述动作信息携带于所述 STA 上报的第一帧的动作详细信息字段中；

所述动作详细信息字段包括时间 1 指示子字段和持续时间 1 子字段、……、时间 K 指示子字段和持续时间 K 子字段，所述 K 为正整数；

所述时间 1 指示子字段用于指示动作 1 的开始时间/结束时间，所述动作 1 是所述 STA 在上一次成功接收第一响应信息后，第一次发送报文且未接收到第二响应信息时的发送动作，所述第一响应信息是所述 AP 成功接收所述 STA 发送的动作信息时发送的响应信息，所述第二响应信息是所述 AP 成功接收所述 STA 发送的报文时发送的响应信息，所述持续时间 1 子字段用于指示所述动作 1 的持续时间；

所述时间 K 指示子字段用于指示动作 K 的开始时间/结束时间，所述动作 K 是所述 STA 在上一次成功接收所述第一响应信息后，第 K 次发送报文且未接收到所述第二响应信息时的发送动作，所述持续时间 K 子字段用于指示所述动作 K 的持续时间。

59. 根据权利要求 51 或 52 所述的装置，其特征在于，所述动作信息携带于所述 STA 上报

的第一帧的动作详细信息字段中；

所述动作详细信息字段包括第一时间 1 指示子字段和第一时间 1 指示子字段、.....、第一时间 K 指示子字段和第一时间 K 指示子字段，所述 K 为正整数；

所述第一时间 1 指示子字段用于指示动作 1 的起始时间，所述动作 1 是所述 STA 在上一次成功接收第一响应信息后，第一次发送报文且未接收到第二响应信息时的发送动作，所述第一响应信息是所述 AP 成功接收所述 STA 发送的动作信息时发送的响应信息，所述第二响应信息是所述 AP 成功接收所述 STA 发送的报文时发送的响应信息，所述第二时间 1 指示子字段用于指示所述动作 1 的结束时间；

所述第一时间 K 指示子字段用于指示动作 K 的起始时间，所述动作 K 是所述 STA 在上一次成功接收所述第一响应信息后，第 K 次发送报文且未接收到所述第二响应信息时的发送动作，所述第二时间 K 指示子字段用于指示所述动作 K 的结束时间。

60. 根据权利要求 53 所述的装置，其特征在于，所述动作信息和所述载波侦听结果信息携带于所述 STA 上报的第一帧的动作详细信息字段中；

所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段，所述 T 为正整数；

所述时间指示子字段用于指示所述 STA 上一次成功接收第一响应信息的时刻，所述第一响应信息是所述 AP 成功接收所述 STA 发送的动作信息时发送的响应信息；

所述数据 1 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息后的第一个时隙内的动作和载波侦听结果；

所述数据 T 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息后的第 T 个时隙内的动作和载波侦听结果。

61. 根据权利要求 53 所述的装置，其特征在于，所述动作信息和所述报文传输结果信息携带于所述 STA 上报的第一帧的动作详细信息字段中；

所述动作详细信息字段包括时间指示子字段、数据 1 子字段至数据 T 子字段，所述 T 为正整数；

所述时间指示子字段用于指示 STA 上一次成功接收第一响应信息的时刻，所述第一响应信息是所述 AP 成功接收所述 STA 发送的动作信息时发送的响应信息；

所述数据 1 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息后的第一个时隙内的动作和报文传输结果；

所述数据 T 子字段用于指示所述 STA 在上一次成功接收所述第一响应信息后的第 T 个时隙内的动作和报文传输结果。

62. 根据权利要求 51 至 61 任一项所述的装置，其特征在于，所述处理单元根据所述第一神经网络的训练结果更新所述第一神经网络，并在侦听到信道空闲时根据更新后的第一神经网络和所述 STA 当前的状态信息确定是否接入信道，具体用于：

将该 STA 当前的状态信息输入到更新后的第一神经网络，输出第一值和第二值，所述第一值用于表征接入信道获得的预期奖励，所述第二值用于表征不接入信道获得的预期奖励；

在所述第一值大于所述第二值时，确定接入信道；

在所述第一值小于所述第二值时，确定不接入信道。

63.一种通信装置，其特征在于，包括处理器和收发器，所述收发器用于与其它通信装置进行通信；所述处理器用于运行程序，以使得所述通信装置实现权利要求 1 至 19 任一项所述的方法，或者，以使得所述通信装置实现权利要求 20 至 31 任一项所述的方法。

64.一种计算机可读存储介质，所述计算机可读存储介质存储有指令，当其在计算机上运行时，使得权利要求 1 至 19 任一项所述的方法被执行；或者权利要求 20 至 31 任一项所述的方法被执行。

65.一种包含指令的计算机程序产品，当其在计算机上运行时，使得权利要求 1 至 19 任一项所述的方法被执行；或者权利要求 20 至 31 任一项所述的方法被执行。

66.一种通信系统，其特征在于，包括权利要求 32 至 50 任一项所述的通信装置，以及权利要求 51 至 62 任一项所述的通信装置。