



专题：5G 演进技术

## 基于人工智能的大规模天线信道状态信息反馈研究

黄秋萍<sup>1,2</sup>, 刘晓峰<sup>3</sup>, 高秋彬<sup>1,2</sup>, 刘正宣<sup>1,2</sup>, 金立强<sup>1,2</sup>, 孙韶辉<sup>1,2</sup>

- (1. 中信科移动通信技术股份有限公司, 北京 100083;  
2. 无线移动通信国家重点实验室(电信科学技术研究院有限公司), 北京 100191;  
3. 中国信息通信研究院, 北京 100191)

**摘要:** 信道状态信息(channel state information, CSI)的精确获取是大规模天线发挥效能的关键。在现有的通信系统中, 上下行链路互易性不理想时, 基于码本进行下行链路的 CSI 反馈。随着天线规模的增大, 码本 CSI 反馈所需要的开销也越来越大。给出了基于人工智能(artificial intelligence, AI)的 CSI 反馈压缩方法, 分析了基于 AI 的 CSI 反馈的标准化影响、通信流程与面临的挑战, 提供了评估结果。评估结果表明, 相对于基于频域基向量压缩的码本 CSI 反馈, 基于 AI 的 CSI 反馈在相同的反馈精度下可以显著地降低反馈开销。

**关键词:** 大规模天线; 人工智能; 信道状态信息反馈

中图分类号: TP393

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2022051

## Study of channel state information feedback based on artificial intelligence

HUANG Qiuping<sup>1,2</sup>, LIU Xiaofeng<sup>3</sup>, GAO Qiubin<sup>1,2</sup>, LIU Zhengxuan<sup>1,2</sup>, JIN Liqiang<sup>1,2</sup>, SUN Shaohui<sup>1,2</sup>

1. CICT Mobile Communications Technology Co., Ltd., Beijing 100083, China  
2. State Key Laboratory of Wireless Mobile Communications, China Academy of Telecommunications Technology (CATT), Beijing 100191, China  
3. China Academy of Information and Communications Technology (CAICT), Beijing 100191, China

**Abstract:** Accurate acquisition of CSI (channel state information) is the key to the performance of massive MIMO. In current communication systems, when the reciprocity of uplink and downlink is not ideal, codebook-based CSI feedback is used for downlink CSI acquisition. With the increase of antenna scale, codebook-based CSI feedback needs more and more overhead. The CSI feedback compression method based on AI (artificial intelligence) was presented, and the standardization impact, communication process and challenges of CSI feedback based on AI were analyzed. Besides, evaluation results were provided. The evaluation results show that compared with codebook-based CSI feedback based on frequency domain basis vector compression, CSI feedback based on AI can significantly reduce the feedback cost at the same feedback accuracy.

**Key words:** massive MIMO, AI, CSI feedback

收稿日期: 2022-02-10; 修回日期: 2022-03-15

基金项目: 国家重点研发计划项目(No.2020YFB1807100)

**Foundation Item:** The National Key Research and Development Project of China (No.2020YFB1807100)

## 0 引言

随着 5G 新空口 (new radio, NR) 系统的广泛商用和部署, 业界对移动通信系统的研究重点已转向 5G-A (5G-Advanced, 5G 演进) 系统。大规模天线技术是提高频谱效率和系统容量, 满足人们日益提高的通信质量需求的重要技术手段, 是 5G 系统和 5G-A 系统的关键技术<sup>[1-3]</sup>。在大规模天线系统中, 信道状态信息 (channel state information, CSI) 的获取与反馈是决定性能的关键。基站需要基于 CSI, 进行波束管理和调度。对于上行链路, 基站可以利用上行探测参考信号进行上行 CSI 的获取。对于下行链路, 在上下行互易性较好时, 基站可以利用上行探测参考信号进行下行 CSI 的获取, 终端仅需要反馈信道质量信息 (channel quality information, CQI), 以反映其受到的干扰和噪声情况, 在上下行互易性不理想时, CSI 的获取只能依赖终端的反馈。

CSI 反馈包括全信道反馈和信道特征反馈两种方式。考虑开销问题, 5G NR 系统并不支持全信道信息的反馈, 仅支持终端反馈用于表征信道特征的信息和信道质量, 基站基于这些信息重构信道。在这种方式下, 用于表征信道特征的信息的选取和精度尤为重要。在 NR 系统中, CSI 可以反馈的信息包括 CQI、预编码矩阵指示 (precoding matrix indicator, PMI)、秩指示 (rank indicator, RI)、信道状态信息参考信号资源指示 (channel state information-reference signal resource indicator, CRI)、同步信号/物理广播信道块资源指示 (synchronize signal/physical broadcasting channel block resource indicator, SSBRI)、层指示 (layer indicator, LI) 等<sup>[4]</sup>。

随着天线规模的增加, CSI 反馈的开销也随之增加, 不利于系统容量提升。如何降低 CSI 反馈的开销是大规模天线系统的一个研究重点和难点。为了降低 CSI 反馈的开销, 一方面是对 CSI 反

馈量进行压缩, 降低单次 CSI 反馈的开销<sup>[5-7]</sup>; 另一方面是降低 CSI 反馈的频次, 从而降低总的开销<sup>[8]</sup>。在降低单次 CSI 反馈量的开销方面, 业界最常用的方法是对 CSI 反馈量进行压缩。其基本原理是利用信道在时域/空域/频域的相关性, 将信道在某些基底上进行压缩降维, 从而降低 CSI 反馈量的开销。在 5G NR 系统的 3GPP Release 16 (Rel-16) 阶段, 基于 Type-II 码本, 利用频域信道的相关性对各子带的反馈系数进行了压缩, 将 CSI 反馈的维度从子带数目的整数倍降低到了频域基向量维度的整数倍<sup>[4]</sup>。

现有的 CSI 反馈量压缩方法的性能严重依赖于对于信道结构的先验假设。当信道在对应的变换基上不具备稀疏性时, 将大大降低 CSI 反馈的精度, 性能无法保障。鉴于人工智能 (artificial intelligence, AI) 在图像处理和自然语言处理等领域的巨大成功, 无线通信与 AI 技术的结合受到了学术界和工业界的极大关注<sup>[9-13]</sup>。国内外学者们积极探索如何利用 AI 技术降低 CSI 反馈的开销<sup>[14-20]</sup>, 第三代合作伙伴计划 (3rd Generation Partnership Project, 3GPP) 开启了基于 AI 的 CSI 反馈课题的研究, 探索未来的标准化方向<sup>[21]</sup>。

## 1 基于 AI 的 CSI 反馈量压缩

5G NR 系统 CSI 反馈量的开销主要是 PMI 的开销。预编码的目的是对发送信息进行匹配于信道的预处理, 以将信息发送到相应的方向。理想的预编码方式是将信道的特征向量作为预编码矩阵。考虑直接将信道特征向量的幅度和相位进行较高精度的量化开销太大, 5G NR 系统通过码本对特征向量进行量化。利用码本对特征向量进行量化相当于对特征向量做线性/非线性变换。最优的变换方式与场景/天线结构等诸多因素有关, 难以设计和求解。AI 基于数据进行训练, 可以习得最优的变换, 更有效地利用信道结构特征, 以更小的开销达到相同的性能。此外, 由于 AI 是数据驱动, 通过 AI 模型对信道特征进行提取, 还可以使得基于 AI 的 CSI 反馈



量不再限于预编码矩阵、传输层数等。本章介绍基于 AI 的 CSI 反馈量压缩方法。

### (1) 基于 AI 的全信道压缩

5G NR 系统的 CSI 反馈没有采用全信道反馈，其主要原因是全信道反馈开销太大。如果 AI 可以大大降低压缩后反馈量的开销，在 5G-A 系统中进行基于 AI 的全信道反馈成为一种可能。

在进行基于 AI 的全信道压缩时，AI 编码器的输入为终端通过信道状态信息参考信号（CSI reference signal, CSI-RS）获得的信道估计值或经过预处理后的信道估计值。基于 AI 的全信道压缩反馈示意图如图 1 所示（图中的虚线框为可选模块）：将终端估计的原始信道或经过预处理后的信道输入 AI 编码器，利用 AI 编码器对信道信息进行压缩后反馈给基站，基站侧执行逆过程恢复信道信息。对终端获得的信道估计值进行预处理的目的是降低 AI 编码器输入的维度，从而降低 AI 训练的复杂度。信道的预处理通常基于信道在时域/空域/频域的稀疏性进行。在一些 AI 译码器的设计中，CSI 反馈的量化值直接作为 AI 译码器的输入，此时不需要解量化的步骤。在 AI 编码器的设计中包括对 CSI 反馈量的量化时，AI 编码器之外不再需要独立的量化模块。

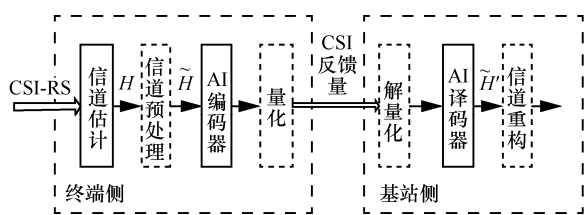


图 1 基于 AI 的全信道压缩反馈示意图

文献[22]提出的基于深度学习的 CSI 反馈架构 CsiNet 应用了图 1 所示的结构。在 CsiNet 中，训练后的 CSI 反馈量的量化执行也在 AI 编码器中。信道的预处理过程如下：终端侧将基于下行参考信号测量得到的频域下行信道进行角度-时延域变换，压缩信道的维度；然后利用信道矩阵在时域信道的稀疏性进行信道裁剪，进一步降低

信道矩阵的维度。预处理后的信道是一个二维信道，若将其维度表示为  $M \times 2$ ，将信道的实部和虚部各使用一个维度表示，则信道数据的维度成为了  $M \times N \times 2$ 。将它看作长度为  $M$ 、宽度为  $N$ 、通道数为 2 的图像。借鉴卷积神经网络进行图像处理的思想，AI 编码器的内部流程为：卷积提取信道特征→重塑→全连接生成压缩量→量化。基站在获取到 CSI 反馈压缩量后，利用对应的 AI 译码器进行信道的重建。CsiNet 的 AI 译码器包括两个 RefineNet 单元、一个卷积层和一个全连接层，两个 RefineNet 的目的是改善重建的信道矩阵。其中，每个 RefineNet 单元中包括多个卷积层。基于残差网络思想，RefineNet 单元中的第一层和最后一层的输出相加作为整个单元的输出，以避免梯度消失问题。为了使得输出值保持原始信道矩阵的维度，RefineNet 单元中的所有卷积层还对末位进行 0 值填充的方法。仿真结果显示，相对于传输的压缩反馈算法，CsiNet 在不同场景的不同压缩率下均有更好的 CSI 重建性能。

### (2) 基于 AI 的信道特征压缩

类似于基于 AI 的全信道反馈，终端可以利用信道估计值进行信道特征提取，然后利用 AI 编码器和译码器进行信道特征的压缩和解压缩。如图 2 所示，终端通过 CSI-RS 进行信道估计，并基于信道估计值提取信道特征  $v$ ，然后 AI 编码器对提取出的信道特征  $v$  进行压缩、量化，然后反馈给基站。基站收到终端的反馈信息后，执行逆过程，即可恢复出 AI 编码器压缩前的信道特征  $v'$ 。在恢复出信道特征  $v'$  后，基站可以进一步地进行信道重构，以恢复出原始信道或输出多用户配对的预编码等。

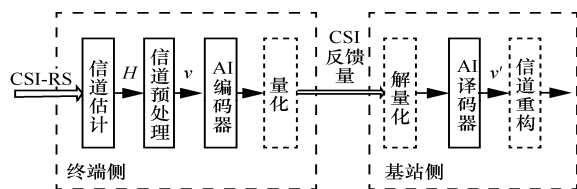


图 2 基于 AI 的信道特征压缩反馈示意图

在图 2 的结构中, 信道特征的提取不参与 AI 编码器的训练。在以反馈预编码为目的时, 终端可以通过对信道估计值进行奇异值分解来获得信道特征向量。除了进行信道特征压缩外, AI 算法也可以用于信道特征提取。此时, 可以将信道特征提取与信道特征压缩放在同一个 AI 编码器里同时训练, 则基于 AI 的信道特征提取和信道特征压缩反馈框图与图 1 相同。

在实际应用中, 可以考虑仅对部分信道特征利用 AI 编码器/译码器进行压缩/解压缩。对部分信道特征利用 AI 模型进行压缩/解压缩的主要应用场景如下。

- 信道特征开销差异较大的场景。在传统的 CSI 反馈方式中部分信道特征开销较大、部分信道特征开销较小时, 可以利用 AI 编码器对开销较大的信道特征进行压缩, 开销较小的信道特征仍然使用传统的反馈方式。例如, 在基于 Type II 码本的 CSI 反馈时, PMI 反馈开销最多可达数百比特, 而 RI 反馈至多需要 3 bit。利用 AI 模型对 PMI 进行压缩反馈具有很大降低开销的潜力, RI 开销的压缩则无增强空间, 没有利用 AI 模型对 RI 进行压缩反馈的必要性。此时可以考虑对 PMI 利用 AI 模型进行压缩反馈, RI 使用传统的方式反馈。
- 信道特征相关性不同的场景。在信道特征的相关性不同时, 可以考虑对时域/频域/空域相关性较强的信道特征利用 AI 模型进行压缩, 降低反馈开销, 相关性较弱的信道特征仍然采用传统的方式进行反馈。这是因为基于数据驱动的 AI 模型可以在信道特征在时域/频域/空域的相关性较强时, 在保证性能的同时大大降低反馈开销。在部分信道特征的上下行相关性较强, 部分信道特征上下行相关性较弱时, 也可以利用 AI 模型对上下行相关性较强的信道

特征进行压缩, 再结合上行信道即可在保证性能的同时大大降低反馈开销。

在仅对部分信道特征进行基于 AI 的 CSI 反馈时, 类似于一个双通道 CSI 反馈的过程, 如图 3 所示。在图 3 的网络结构中, 信道特征  $v_1$  利用 AI 编码器进行压缩, 压缩后通过第一个 CSI 反馈通道反馈, 信道特征  $v_2$  被直接进行量化, 通过第二个 CSI 反馈通道反馈。基站分别恢复  $v_1'$  和  $v_2'$  后, 进行信道的重构。在实际系统中, 这两个 CSI 反馈通道可以通过同一个物理信道联合反馈, 可以在不同时刻分别反馈。

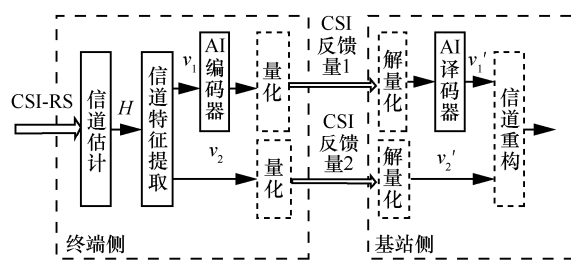


图 3 基于 AI 的双通道信道特征压缩反馈示意图

文献[23]提出的基于上下行链路互易性的 CSI 反馈架构 DualNet-MAG 和 DualNet-ABS 采用了图 3 的基本结构。文献[23]的研究表明, 对于频分双工 (frequency division duplex, FDD) 大规模天线系统, 在将时延域的 CSI 变换为极坐标形式时, 上下行链路的相位相关性不明显, 但幅度具有较强的相关性; 分别观察 CSI 的实部和虚部, 上下行链路 CSI 的符号相关性不明显, 但实部和虚部的绝对值具有较强的相关性。由于上下行链路相关性较强的信道特征可以基于上行链路的信道特征进行调整, 因此将其压缩至较少的反馈量不会影响性能。DualNet-MAG 和 DualNet-ABS 基于该思想, 将上下行链路相关性较强的信道特征进行了压缩, 将上下行链路相关性不强的信道特征直接量化反馈。其中, DualNet-MAG 将信道特征分离为幅度部分和相位部分, 幅度部分通过 AI 编码器进行压缩反馈, 相位部分基于幅度分布量化后进行反



馈。在基站利用 AI 译码器进行幅度部分的恢复时，将上行链路的幅度信息作为输入量输入 AI 译码器，辅助进行幅度部分的恢复。DualNet-ABS 与 DualNet-MAG 的原理类似，区别在于将信道特征分离为实部和虚部各自的绝对值和符号，绝对值部分通过 AI 编码器进行压缩反馈，相位部分基于绝对值分布量化后进行反馈。评估结果表明，在多种应用场景下，DualNet-MAG 和 DualNet-ABS 的性能都优于 CsiNet。在相同的压缩量下，DualNet-MAG 和 DualNet-ABS 相较于 CsiNet 也表现出更强的鲁棒性。

### (3) 基于 AI 的多压缩比 CSI 压缩

无线信道在时间上具有相关性。在信道的相干时间内，利用不同时刻信道间的相关性，可以对部分时刻的 CSI 进行高压缩比（compression ratio, CR）压缩，部分时刻的 CSI 进行低 CR 压缩。其中，CR 的定义为压缩后比特数与压缩前比特数的比值。类似地，利用无线信道在频域上的相关性，也可以对部分频域资源的 CSI 进行高 CR 压缩，部分频域资源的 CSI 进行低 CR 压缩。

文献[24]基于文献[22]的 CsiNet 进行了改进，提出了一种 CsiNet-LSTM CSI 反馈结构。在该结

构中，在基站对信道恢复的过程中增加长短期记忆（long short-term memory, LSTM）模块，并在时间上将待反馈的信道划分成了若干组。在各个组中，仅对最早时刻的信道利用高 CR 进行基于 AI 的压缩/解压缩，以获得较高精度的信道恢复；对于其他时刻的信道，利用低 CR 进行基于 AI 的压缩/解压缩，经过 AI 解压缩后进入 LSTM 网络。每个低 CR AI 译码器的输入为高 CR AI 编码器输出的 CSI 反馈量与当前时刻低 CR AI 编码器输出的 CSI 反馈量的级联。LSTM 网络中包括一些存储单元，用来存储之前信道提取出的信息，以用于后续的预测。LSTM 网络可以从之前提取出的信道信息中学习时间的相关性，结合当前时刻 AI 译码器输出的信道信息，恢复原始信道信息。其原理示意图如图 4 所示。相对于未使用 LSTM 网络且所有时刻的 CSI 反馈使用高 CR AI 编码器进行压缩的方案，CsiNet-LSTM 方案可以在相同信道恢复精度下大大降低总的反馈开销。

文献[25]利用了无线信道在时间上的相关性，基于 CsiNet 提出了 RecCsiNet 和 PR-RecCsiNet CSI 反馈架构。RecCsiNet 和 PR-RecCsiNet 与 CsiNet 的网络结构相同，但 AI 编码器和 AI 译码

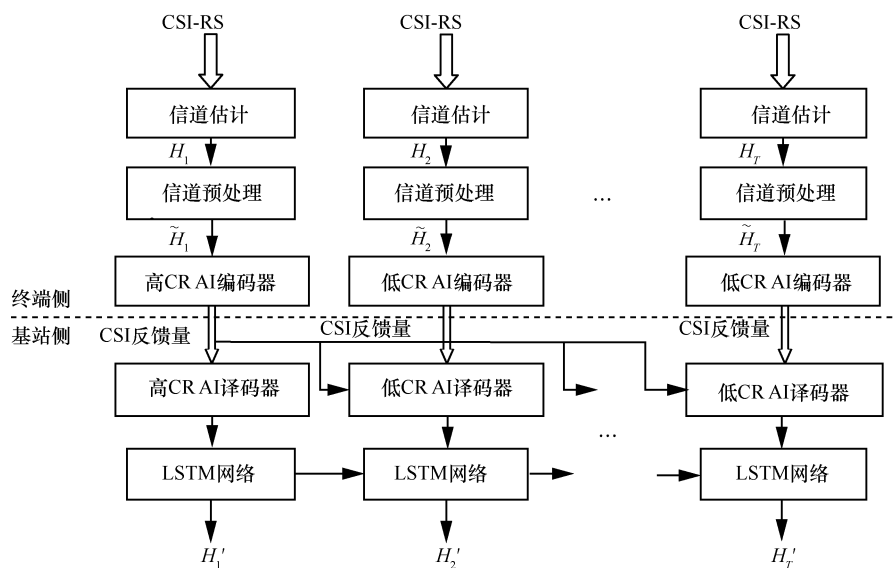


图 4 CsiNet-LSTM 网络示意图

器内部结构不同。在 RecCsiNet 的 AI 编码器中, 首先对信道进行信道特征提取, 然后经过并联的线性全连接网络(fully-connected network, FCN)和 LSTM 网络。CsiNet-LSTM 中只有基站侧使用 LSTM 网络进行信道恢复, RecCsiNet 中的 AI 编码器和 AI 译码器中均有 LSTM 网络。因此, 二者中的 LSTM 网络并不相同。RecCsiNet 的 LSTM 网络中也有存储单元, 且具有提取时变信道的时域相关性的能力。但 LSTM 网络并不用于学习时间相关性, 而是用来学习残差特性, 从而使得信道特征的压缩具有更好的鲁棒性。FCN 被用来作为跳跃连接, 可以加速收敛并减少梯度消失问题。PR-RecCsiNet 与 RecCsiNet 原理类似, 区别在于 LSTM 网络与 FCN 是串联的关系。评估结果表明, 相对于 CsiNet, RecCsiNet 具有更高的信道重建精度。

## 2 标准化影响

基于 AI 的 CSI 反馈对标准化的影响涉及多个方面。

对于基于 AI 的 CSI 反馈量的压缩, 基站侧的 AI 译码器与终端侧的 AI 编码器必须匹配才能保证性能。因此, 收/发两端使用的 AI 模型应当联合训练, 以形成一个完整的 AI 网络。

从标准化的角度看, 基站和终端首先需要交互 AI 模型的相关信息, 以使得收/发两端的 AI 模型相匹配。标准化 AI 模型的结构和参数并不利于设备实现的灵活性。因此, 应当允许训练 AI 编码器/译码器的设备自行设计 AI 模型, 然后将其发送给对侧。AI 模型的相关信息可以包括 AI 模型的基础结构与主要参数, 也可以是一个“黑盒子”, 例如, 为一个可执行程序。当 AI 模型的相关信息为 AI 模型的基础结构与主要参数时, 对侧需要基于这些信息构建对应的 AI 模型。当 AI 模型的相关信息为一个“黑盒子”时, 对侧并不需要了解 AI 模型具体是什么, 只需要输入符合预先约定格式的输入量, 即可输出所需要的信息。因此, 除

了 AI 模型相关信息的传输方式、比特大小和格式外, 还有可能需要标准化 AI 编码器输入量的内容和格式、AI 译码器输出量的内容和格式等。

其次, 需要标准化 CSI 反馈量的传输方式、比特大小和格式。AI 模型的设计不同, AI 编码后获得的 CSI 压缩反馈量可能也有所不同。由于 AI 编码器和 AI 译码器之间交互的反馈量可以认为是 AI 网络内部的信息, 因此无须标准化该反馈量表征的具体内容是什么, 只需要标准化该反馈量的比特大小和格式, 使得基站可以解析该反馈量对应的比特。由于基于 AI 的 CSI 反馈压缩下 CSI 反馈量反馈的比特大小和格式与 5G NR 系统当前基于码本的 CSI 反馈不同, 因此不排除需要标准化新的 CSI 反馈方式。

此外, 训练 AI 模型的设备不同, 标准化的影响也可能不同。理论上, AI 模型的训练可以在终端侧进行, 也可以在基站侧进行。CSI 反馈需要压缩的是下行信道或下行信道特征。由于下行信道只能由终端通过下行参考信号测量获得, 因此终端侧更易于利用实测数据进行 AI 模型训练。但终端的存储空间和电池容量有限, 数据集的存储势必会占用终端的存储空间, 模型的训练对终端的功耗是一大挑战。此外, 终端基于测量数据训练的模型通常具有局限性, 难以适用于其他终端。如果模型的训练在终端侧进行, 基站需要与各终端分别交互 AI 模型的相关信息, 这对上行链路来说是一个不小的挑战。因此, 在基站侧进行 AI 模型的训练更为合理。基站可以收集多个终端的信道信息构建数据集, 训练出一套适用于多个终端的 AI 模型, 然后将该模型的相关信息发送给各终端。对于共用同一个 AI 模型的终端, 基站可以通过广播/多播的方式向终端发送 AI 模型相关信息。当 AI 模型的训练在基站侧进行时, 终端需要向基站反馈信道特征/信道的原始数据, 以用于数据集的构建和模型训练。终端如何反馈信道特征/信道的原始数据, 也是标准化需要考虑的内容。信道特征/信道的原始数据除了可以用于数据集



的初始构建和 AI 模型的初始训练外，还可以用于 AI 模型的更新。基站在对 AI 模型进行更新后，需要向终端发送 AI 模型的更新信息，以保证收发两端 AI 模型的匹配。

基于上述分析，一种潜在的基于 AI 的 CSI 压缩反馈流程示意图如图 5 所示。

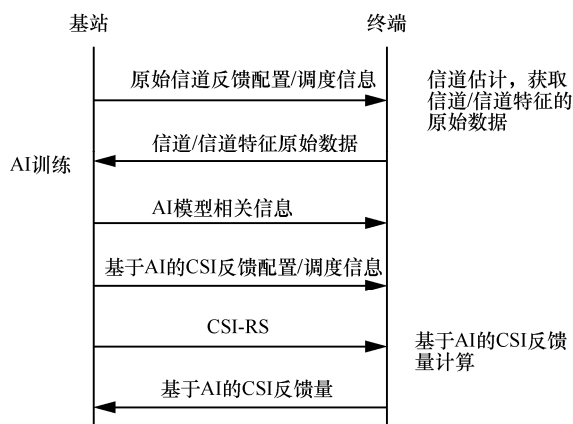


图 5 基于 AI 的 CSI 压缩反馈流程示意图

### 3 研究挑战与机会

虽然大量的学术研究证明了基于 AI 的 CSI 反馈的性能增益，将 AI 用于实际系统仍然面临着大量的挑战。

#### (1) 网络架构的创新

现有的基于 AI 的 CSI 反馈网络结构多是终端使用 AI 编码器进行 CSI 反馈量的压缩，基站使用 AI 译码器进行 CSI 反馈量的恢复。在这种网络架构下，为了保证 AI 编码/译码的性能，需要对 AI 编码器和译码器联合训练。这要求 AI 编码器和译码器在同一侧设备进行训练，然后将 AI 编码器/译码器的信息传递给另一侧。一方面，AI 编码器/译码器信息的传递带来了 CSI 反馈量之外的额外开销；另一方面，AI 编码器的训练需要关于信道的原始数据。如果在基站侧进行训练，如何获取信道的原始数据是一个需要解决的问题；如果在终端侧进行训练，AI 训练会带来增加终端复杂度和功耗等问题。为了解决上述问题，需要探索无须收/

发两侧交互 AI 模型信息的方案。例如，探索收/发两侧独立进行 AI 训练的方案，或者，探索在单侧应用 AI 模块，另一侧提供少量信息辅助的方案。

#### (2) 数据集的获取

AI 模型的训练离不开数据集。模拟的数据无法体现实际的信道性能，难以根据实际的信道条件更新模型参数。如何基于实际的数据进行 AI 模型的训练，如何进行训练所需数据集的获取是实际应用时需要考虑的问题。

#### (3) AI 训练的复杂度

AI 模型的好坏与数据集的大小息息相关。过小的数据集训练的 AI 模型鲁棒性较差，也无法保证 CSI 反馈的精度。较大规模的数据集可以保证 AI 模型的性能，但训练复杂度也随之增加。如何降低 AI 训练的复杂度是实际应用的一大难点。此外，现有的研究中 AI 模型多是静态训练的。静态训练的 AI 模型难以匹配信道的动态变化，如何以较低的复杂度实现 AI 模型的动态更新也是需要考虑的问题。

#### (4) AI 模型的泛化问题

AI 模型的泛化包括两个层面的内容：一方面是用户间的模型泛化，另一方面是同一个用户在多种信道条件下的模型泛化。前者是为了降低模型训练的复杂度，降低关于 AI 模型信息的交互量；后者是为了提高 AI 模型的鲁棒性。当前的研究多集中在单用户基于特定数据集的模型训练。为了在实际系统中进行更好的应用，AI 模型的泛化是一个非常重要的课题。

#### (5) 从理论到实际的非理想因素

现有的基于 AI 的 CSI 反馈研究中所提供的性能多假设收/发两端已知训练好的 AI 模型，并没有考虑 AI 模型相关信息的交互是否理想，以及 AI 模型相关信息交互的开销。在实际应用中，需要考虑信息交互中的不理想因素。

## 4 性能评估

本节评估了图 2 所示的基于 AI 的信道特征压

缩反馈的性能。仿真假设见表 1, 其中,  $M$ 、 $N$ 、 $P$ 、 $M_g$ 、 $N_g$ 、 $M_p$ 、 $N_p$  分别表示垂直方向天线阵子数、水平方向天线阵子数、天线极化数、垂直方向天线面板数、水平方向天线面板数、垂直方向天线阵子映射后的发送端口数和水平方向天线阵子映射后的发送端口数,  $dH$  和  $dV$  分别表示水平方向的阵子间距和垂直方向阵子间距。评估所使用的 AI 网络如图 6 所示。AI 网络的输入为基于信道估计获取的信道特征向量, 所使用的 AI 编码器和 AI 译码器的结构如下: 编码器包含两个 TB 块(transformer based block)、一个全连接层(Dense 层)以及量化器, TB 块用于提取子带内和子带之间 CSI 特征, 全连接层用于下采样压缩 CSI, 尾部级联的 Sigmoid 激活函数将输出归一化到[0,1]便于量化器量化, 量化器将输入的实数量化为二进制比特数; 类似地, 译码器也包含两个 TB 块和一个全连接层以及解量化器, 解量化器将输入的二进制比特数映射为[0,1]取值范围的实数, 经过一个全连接层上采样相匹配的维度, 再经过两个 TB 块恢复重构 CSI。其中, TB 块采用了类似

Transformer 模型的多头注意力机制、残差和层归一化结构。量化器采用了[0,1]上均匀的 2 比特量化。通过调节编码器的全连接层输出维度, 可以实现不同反馈开销。在图 7 的评估结果中, 同时提供了采用 5G NR 系统基于频域基向量压缩的增强 Type II 码本进行 CSI 反馈时的性能作为比较基准, 比较两种方式下基站恢复的信道特征向量与实际的信道特征向量的余弦相似度。余弦相似度的计算式如式(1)所示, 用来表征两个向量之间的相关性。

$$\rho = \frac{|\mathbf{b}^H \mathbf{v}|}{\|\mathbf{b}_2\| \|\mathbf{v}_2\|} \quad (1)$$

其中,  $\mathbf{v}$  表示信道特征向量; 对于增强 Type II 码本,  $\mathbf{b}$  表示终端反馈的预编码矩阵; 对于基于 AI 的 CSI 反馈,  $\mathbf{b}$  表示 AI 译码器恢复出的信道特征向量。

根据图 7 的评估结果可以看出, 在相同的余弦相似度下, 基于 AI 的信道特征压缩反馈相对于增强 Type II 码本 CSI 反馈可以节省 80~150 bit 的开销。

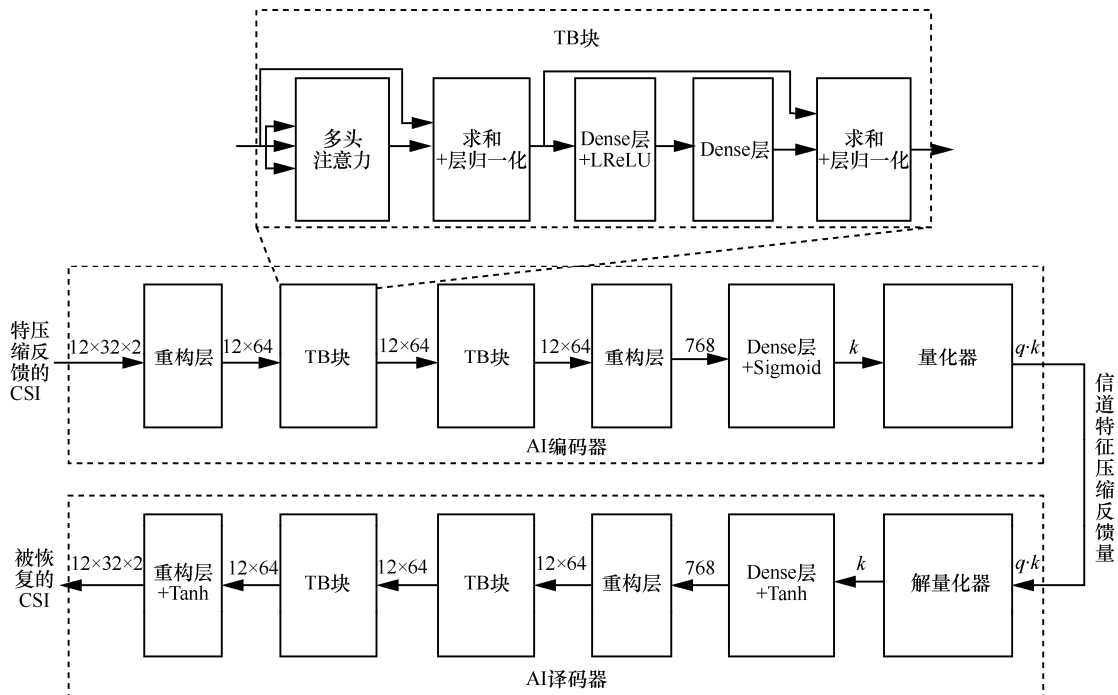


图 6 基于 AI 的信道特征压缩反馈 AI 网络结构



表 1 仿真假设

| 参数名称      | 值                                                                                               |
|-----------|-------------------------------------------------------------------------------------------------|
| 双工方式/波形   | FDD/OFDM                                                                                        |
| 场景        | 密集城区                                                                                            |
| 频点范围      | 2 GHz, 上下行载频间隔为 200 MHz                                                                         |
| 站间距       | 200 m                                                                                           |
| gNB 的天线配置 | 32 端口: $(M, N, P, M_g, N_g, M_p, N_p) = (8, 8, 2, 1, 1, 2, 8)$ , $(dH, dV) = (0.5, 0.8)\lambda$ |
| UE 的天线配置  | 4 端口: $(M, N, P, M_g, N_g, M_p, N_p) = (1, 2, 2, 1, 1, 1, 2)$ , $(dH, dV) = (0.5, 0.5)\lambda$  |
| 基站发送功率    | 49 dBm                                                                                          |
| 基站天线高度    | 29 m                                                                                            |
| 系统带宽      | 10 MHz                                                                                          |
| 子载波间隔     | 15 kHz                                                                                          |
| CQI 子带    | 4 PRB                                                                                           |
| 传输层数      | 最大 12 层                                                                                         |
| CSI 反馈    | CSI 反馈周期: 5 ms; 调度时延: 4 ms                                                                      |
| 业务类型      | FTP 模式 1, 包大小为 0.5 MB                                                                           |
| 资源利用率     | 70%                                                                                             |
| UE 分布     | 室内 80% (3 km/h), 室外 20% (30 km/h)                                                               |
| UE 接收算法   | MMSE-IRC                                                                                        |

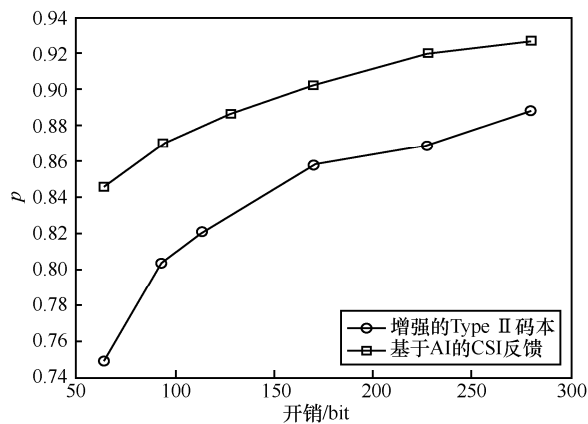


图 7 基于 AI 的 CSI 反馈和增强 Type II 码本 CSI 反馈的平均余弦相似度的比较

## 5 结束语

本文分析了大规模天线系统中的 CSI 反馈量压缩问题,总结了基于 AI 全信道压缩反馈方法、基于 AI 的信道特征压缩反馈方法和基于 AI 的多压缩比 CSI 反馈方法,并进行了基于 AI 的 CSI 反馈的标准化探析,指出了在 5G-A 系统中进行基于 AI 的 CSI 反馈的研究方向。评估结果表明,在相同的余弦相似度下,相对于传统的基于码本

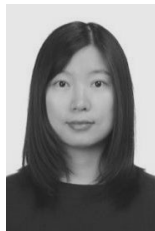
的 CSI 反馈,基于 AI 的信道特征反馈方法可以显著降低 CSI 反馈的开销。

## 参考文献:

- [1] CHATAUT R, AKL R. Massive MIMO systems for 5G and beyond networks-overview, recent trends, challenges, and future research direction[J]. Sensors (Basel, Switzerland), 2020, 20(10): 2753.
- [2] BJORNSEN E, VAN DER PERRE L, BUZZI S, et al. Massive MIMO in sub-6 GHz and mmWave: physical, practical, and use-case differences[J]. IEEE Wireless Communications, 2019, 26(2): 100-108.
- [3] 陈山枝, 孙韶辉, 苏昕, 等. 大规模天线波束赋形技术原理与设计[M]. 北京: 人民邮电出版社, 2019.  
CHEN S Z, SUN S H, SU X, et al. Principles and design of massive beamforming technology[M]. Beijing: Posts & Telecom Press, 2019.
- [4] 3GPP. NR; Physical layer procedures for data (V17.0.0): TS 38.214 [S]. 2021.
- [5] KUO P H, KUNG H T, TING P G. Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays[C]//Proceedings of 2012 IEEE Wireless Communications and Networking Conference. Piscataway: IEEE Press, 2012: 492-497.
- [6] SIM M S, PARK J, CHAE C B, et al. Compressed channel feedback for correlated massive MIMO systems[J]. Journal of Communications and Networks, 2016, 18(1): 95-104.
- [7] GAO Z, DAI L L, WANG Z C, et al. Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO[J]. IEEE Transactions on Signal Processing, 2015, 63(23): 6169-6183.
- [8] CHEON H, PARK B, HONG D. Adaptive multicarrier system with reduced feedback information in wideband radio channels[C]//Proceedings of Gateway to 21st Century Communications Village, IEEE VTS 50th Vehicular Technology Conference (Cat. No.99CH36324). Piscataway: IEEE Press, 1999: 2880-2884.
- [9] YE H, LI G Y, JUANG B H. Power of deep learning for channel estimation and signal detection in OFDM systems[J]. IEEE Wireless Communications Letters, 2018, 7(1): 114-117.
- [10] YANG Y W, GAO F F, LI G Y, et al. Deep learning-based downlink channel prediction for FDD massive MIMO system[J]. IEEE Communications Letters, 2019, 23(11): 1994-1998.
- [11] KANG J M, CHUN C J, KIM I M. Deep-learning-based channel estimation for wireless energy transfer[J]. IEEE Communications Letters, 2018, 22(11): 2310-2313.
- [12] GUO J J, WEN C K, JIN S, et al. Convolutional neural net-

- work-based multiple-rate compressive sensing for massive MIMO CSI feedback: design, simulation, and analysis[J]. IEEE Transactions on Wireless Communications, 2020, 19(4): 2827-2840.
- [13] HSIEH C H, CHEN J Y, NIEN B H. Deep learning-based indoor localization using received signal strength and channel state information[J]. IEEE Access, 2019(17): 33256-33267.
- [14] WANG T Q, WEN C K, WANG H Q, et al. Deep learning for wireless physical layer: opportunities and challenges[J]. China Communications, 2017, 14(11): 92-111.
- [15] LU Z L, WANG J T, SONG J. Multi-resolution CSI feedback with deep learning in massive MIMO system[C]//Proceedings of ICC 2020 - 2020 IEEE International Conference on Communications. Piscataway: IEEE Press, 2020: 1-6.
- [16] LU Z L, WANG J T, SONG J. Binary neural network aided CSI feedback in massive MIMO system[J]. IEEE Wireless Communications Letters, 2021, 10(6): 1305-1308.
- [17] ZHANG Y Y, ZHANG X C, LIU Y. Deep learning based CSI compression and quantization with high compression ratios in FDD massive MIMO systems[J]. IEEE Wireless Communications Letters, 2021, 10(10): 2101-2105.
- [18] LIU Z Y, DEL ROSARIO M, DING Z. A Markovian model-driven deep learning framework for massive MIMO CSI feedback[J]. IEEE Transactions on Wireless Communications, 2022, 21(2): 1214-1228.
- [19] CAO Z, SHIH W T, GUO J J, et al. Lightweight convolutional neural networks for CSI feedback in massive MIMO[J]. IEEE Communications Letters, 2021, 25(8): 2624-2628.
- [20] GAO M, LIAO T M, LU Y B. Fully connected feed forward neural networks based CSI feedback algorithm[J]. China Communications, 2021, 18(1): 43-48.
- [21] 3GPP. RP-213560: New SI: study on artificial intelligence (AI)/machine learning (ML) for NR air interface[S]. 2021.
- [22] WEN C K, SHIH W T, JIN S. Deep learning for massive MIMO CSI feedback[J]. IEEE Wireless Communications Letters, 2018, 7(5): 748-751.
- [23] LIU Z Y, ZHANG L, DING Z. Exploiting Bi-directional channel reciprocity in deep learning for low rate massive MIMO CSI feedback[J]. IEEE Wireless Communications Letters, 2019, 8(3): 889-892.
- [24] WANG T Q, WEN C K, JIN S, et al. Deep learning-based CSI feedback approach for time-varying massive MIMO channels[J]. IEEE Wireless Communications Letters, 2019, 8(2): 416-419.
- [25] LU C, XU W, SHEN H, et al. MIMO channel information feedback using deep recurrent network[J]. IEEE Communications Letters, 2019, 23(1): 188-191.

## [作者简介]



**黄秋萍** (1987- ), 女, 博士, 中信科移动通信技术股份有限公司高级工程师, 主要研究方向为大规模天线技术、移动通信新技术研究与标准制定等。



**刘晓峰** (1981- ), 男, 博士, 中国信息通信研究院无线通信创新中心副总工程师、正高级工程师, 主要研究方向为移动通信系统设计等。



**高秋彬** (1980- ), 男, 博士, 中信科移动通信技术股份有限公司正高级工程师, 主要研究方向为通信系统协议设计、多天线系统、协作传输、信号处理算法以及系统建模与评估等。



**刘正宣** (1982- ), 男, 博士, 中信科移动通信技术股份有限公司工程师, 主要研究方向为大规模多输入多输出和基于人工智能的信道状态信息反馈等。



**金立强** (1991- ), 男, 博士, 无线移动通信国家重点实验室(电信科学技术研究院有限公司)博士后研究员, 主要研究方向为信道编码、AI 在通信物理层应用等。



**孙韶辉** (1972- ), 男, 博士, 中信科移动通信技术股份有限公司副总经理、教授级高级工程师, 主要研究方向为移动通信系统设计与多天线技术、卫星通信和定位等关键技术。