

基于 AAT 模型的毫米波大规模 MIMO 系统信道估计

于舒娟, 刘荣, 张昀, 谢娜, 黄丽亚

(南京邮电大学电子与光学工程学院、柔性电子(未来技术学院), 江苏 南京 210023)

摘要: 针对毫米波大规模多输入多输出信道具有时间相关性、系统易受噪声因素影响导致信道估计精度低的问题, 提出了一种基于改进的时序卷积神经网络信道估计方法。该方法将仿真获得的信道矩阵以二维图像数据方式输入系统; 利用时间相关性进行特征融合, 构建集中注意力机制网络, 增强系统模型对信道深层特征的提取能力; 将 AAN 嵌入时序卷积神经网络中进行训练; 系统输出去噪后的二维图像, 即信道估计矩阵。仿真结果表明, 所提信道估计方法在性能和复杂度方面优于传统的信道估计方法, 并且当测试场景发生改变时依旧具有鲁棒性。

关键词: 大规模多输入多输出信道; 时序卷积神经网络; 信道估计; 集中注意力机制网络

中图分类号: TN92

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2024067

AAT model based channel estimation for mmWave massive MIMO systems

YU Shujuan, LIU Rong, ZHANG Yun, XIE Na, HUANG Liya

College of Electronic and Optical Engineering & College of Flexible Electronics (Future Technology),
Nanjing University of Posts and Telecommunications, Nanjing 210023, China

Abstract: To solve the problems of temporal correlation and susceptibility to noise in millimeter wave massive MIMO channels, which result in decreased channel estimation accuracy, a novel channel estimation method based on an improved temporal convolutional network was proposed. The channel matrices obtained from simulation were feed into the system as two-dimensional image data. The temporal correlation was utilized for feature fusion and an attention in attention network was constructed to enhance the system's ability to extract deep channel features. Then, AAN was integrated into the temporal convolutional network for training. Finally, the system outputted a denoised two-dimensional image, namely, the channel estimation matrix. Simulation results demonstrate that the proposed method not only exhibits good performance and complexity compared to conventional channel estimation methods but also maintains robustness when the test scenario changes.

Keywords: massive MIMO channel, temporal convolutional neural network, channel estimation, attention in attention network

0 引言

5G 的超高速率、超低时延、超大连接等特点为移动终端带来更快的传输速率, 进一步提高了用户使用体验^[1]。毫米波大规模多输入多输出(MIMO, multiple input multiple output) 技术是 5G 系统的关键技术之一, 具有系统稳定性强、频谱利用率高、

系统容量大等特点。在毫米波大规模 MIMO 系统中, 为降低能量消耗和硬件成本, 通常采用数模混合处理结构^[2], 将少量射频(RF, radio frequency)链通过相移网络连接到大量天线上^[3]。通过准确估计信道状态信息, 利用多天线的优势提升自由度, 进而提高信道容量^[4]。

毫米波大规模 MIMO 系统的传统信道估计方

收稿日期: 2023-12-20; 修回日期: 2024-03-07

通信作者: 刘荣, 1311108320@qq.com

基金项目: 国家自然科学基金资助项目(No.61977039)

Foundation Item: The National Natural Science Foundation of China (No.61977039)

法主要有最小二乘 (LS, least square)、最小均方误差 (MMSE, minimum mean square error) 和压缩感知 (CS, compressed sensing) 算法。文献[5]提出了一种用于 MIMO 正交频分复用系统的迭代 LS 信道估计算法。文献[6]采用压缩感知技术来减少信道估计训练和反馈过程中的资源开销, 并提出了一种分布式压缩感知信道估计方法, 该方法通过用户端获取的压缩信号使基站端完成信道的准确估计。为了避免直接处理大量 MIMO 信道矩阵, 有效提高信道估计性能, Fan 等^[7]利用离散傅里叶变换后的信道峰值计算到达方向, 然后利用 LS 算法估计信道增益并计算信道矩阵。文献[8]提出了一种利用毫米波信道稀疏特性的新型毫米波 MIMO 信道估计方法, 该方法使用稀疏贝叶斯学习 (SBL, sparse Bayesian learning) 来粗略估计波束空间域中的信道参数, 然后重建通道, 能够精确地确定多路径成分的角度、时延和增益。

上述传统的信道估计方法在更实际信道模型中性能会有所降低, 并且具有较高的复杂度。相比之下, 基于深度学习的信道估计方法可以利用大量的训练数据提取和处理隐藏特征, 并通过高效的并行计算方法以更低的复杂度更准确地估计信道。文献[9]采用传统的 LS 算法进行信道初步估计, 然后作为图像信息输入深度卷积神经网络 (CNN, convolutional neural network) 中, 提取信道的空间和频率特征信息, 进而进行优化的信道估计。文献[10]中使用长短期记忆 (LSTM, long and short term memory) 网络学习 MIMO 信道训练样本中的时间相关性, 在信道的估计精度上取得了一定的优势。文献[11]开发了一种条件生成对抗性网络, 通过对抗性训练 2 个深度学习网络来预测更真实的信道。文献[12]提出了简化学习近似消息传递网络, 利用预先学习的先验参数集和 Wasserstein 生成对抗网络来提高训练的稳定性效率。通常情况下, 神经

网络深度的增加会提高模型的拟合能力, 但随着网络深度的增加, 也会使模型出现性能退化、复杂度高、过拟合等问题。

针对上述毫米波 MIMO 系统信道估计方法在大规模多输入多输出模型下估计性能的不足, 本文提出了一种基于 AAT (attention in attention temporal) 卷积神经网络模型 (简称 AAT 模型) 的信道估计方法, 主要贡献总结如下。

1) 给出适用于通信信道估计的轻型集中注意力机制网络 (AAN, attention in attention network)。以增强信道的特征信息为目标, 优化当前图像处理超分辨率网络中的最新轻量型集中注意力机制网络, 构建应用于毫米波 MIMO 信道估计的 AAN, 依靠动态注意力分配模块的灵活性, 精准提取相邻子载波信道间的深层特征, 降低无效特征对估计性能的影响。

2) 将 AAN 嵌入时序卷积神经 (TCN, temporal convolutional neural) 网络中, 提出 AAT 模型。考虑到毫米波 MIMO 信道的时间相关性, 本文选用融合时序卷积神经网络提升对信道数据前后向的感知能力, 利用 AAN 根据不同的输入特征对注意力进行合理的动态分配, 充分挖掘数据集中所包含的重要信息。

3) 提出基于 AAT 模型的毫米波信道估计方法。将模拟仿真获得的信道矩阵以二维图像数据方式输入毫米波 MIMO 系统进行去噪处理, 通过 TCN 提取信道特征, 结合 AAN 增强图像的特征信息并进行特征融合提高图像精度。

1 毫米波大规模 MIMO 系统模型

本文考虑一个典型的毫米波大规模多输入多输出场景, 系统模型如图 1 所示^[13]。其中, 发射端部署 N_T 根发送天线和 N_T^{RF} 条 RF 链, 接收端部署 N_R 根接收天线和 N_R^{RF} 条 RF 链, 且 $N_T^{RF} \ll N_T, N_R^{RF} \ll N_R$ 。

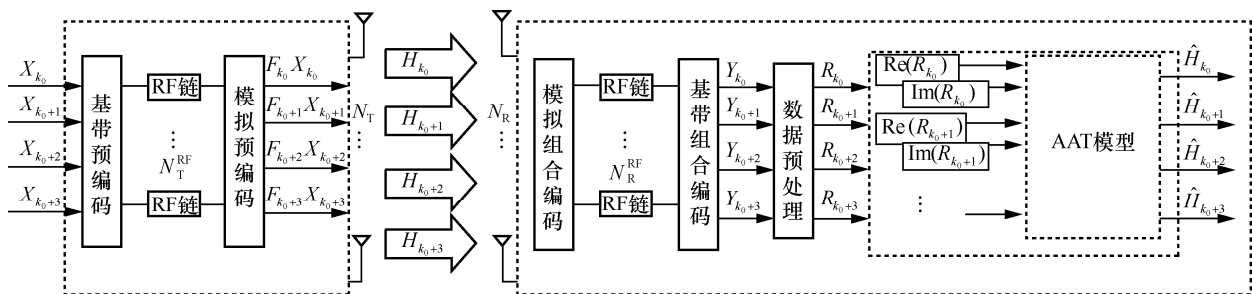


图 1 毫米波大规模 MIMO 系统模型

接收端和发射端之间的信道矩阵可以表示为^[14]

$$\mathbf{H}(\tau) = \sqrt{\frac{N_T N_R}{L}} \sum_{l=1}^L \alpha_l \delta(\tau - \tau_l) \mathbf{a}_R(\varphi_l) \mathbf{a}_T^H(\phi_l) \quad (1)$$

其中, L 为主路径数, $\alpha_l \sim \mathcal{CN}(0, \sigma_\alpha^2)$ 为第 l 条路径的传播增益, σ_α^2 为平均增益功率, τ_l 为第 l 条路径的时延, φ_l 和 $\phi_l \in [0, 2\pi]$ 分别为接收机和发射机到达和离开的方位角 (AoA 和 AoD)。对于均匀线性阵列 (ULA, uniform linear array), 发射端和接收端的阵列导向矢量分别表示为

$$\mathbf{a}_T(\phi_l) = \frac{1}{\sqrt{N_T}} \left[1, e^{-j2\pi \frac{d}{\lambda} \sin(\phi_l)}, \dots, e^{-j2\pi \frac{d}{\lambda} (N_T-1) \sin(\phi_l)} \right]^T \quad (2)$$

$$\mathbf{a}_R(\varphi_l) = \frac{1}{\sqrt{N_R}} \left[1, e^{-j2\pi \frac{d}{\lambda} \sin(\varphi_l)}, \dots, e^{-j2\pi \frac{d}{\lambda} (N_R-1) \sin(\varphi_l)} \right]^T \quad (3)$$

其中, d 和 λ 分别表示相邻天线之间的距离与载波波长。对于时变信道, 在正交频分复用 (OFDM, orthogonal frequency division multiplexing) 中, 第 k_0 个子载波的频域信道可以表示为^[15]

$$\mathbf{H}_{k_0}(t) = \sqrt{\frac{N_T N_R}{L}} \sum_{l=1}^L \alpha_l e^{-j2\pi \left(\tau_l f_s \frac{k_0}{K} - \nu_l t \right)} \mathbf{a}_R(\varphi_l) \mathbf{a}_T^H(\phi_l) \quad (4)$$

其中, ν_l 表示第 l 条路径的多普勒频移, K 表示 OFDM 的子载波数。

虽然在发射机上同时激活具有不同波束的多个射频频链可以加快导频传输过程, 但不能捕获主路径并导致较差的信道估计性能。因此, 发射机仅使用一条射频频链, 向一个波束方向发射导频, 接收机则使用不同波束方向的所有射频频链, 组合接收导频。在不失一般性的情况下, 本文假设发送的原始信号是 \mathbf{X}_{k_0} , 则在接收机基带中, 接收的第 k_0 个子载波的导频信号可以表示为

$$\mathbf{Y}_{k_0} = \mathbf{W}_{k_0}^H \mathbf{H}_{k_0} \mathbf{F}_{k_0} \mathbf{X}_{k_0} + \mathbf{W}_{k_0}^H \mathbf{N}_{k_0} \quad (5)$$

其中, \mathbf{N}_{k_0} 表示组合前的加性白高斯噪声, \mathbf{F}_{k_0} 表示发射机预编码, \mathbf{W}_{k_0} 表示接收机组合编码。

在时域和频域都考虑导频的插入, 具体来说, 相邻的 Q 个子载波分别在一个相干区间的起点放置长度相同的导频, 形成一个导频子载波块, 每个相干区间的其余时隙用于数据传输。2 个导频子载

波块由专用于数据传输的 $Q_d (Q_d \geq 0)$ 个子载波分隔。利用导频估计相应时频位置的信道。基于所估计的信道, 可以应用插值来获得没有导频处的信道。很明显, 插值的精度取决于估计信道的精度和信道的变化。因此, 本文将重点提高基于导频的信道估计的精度, 以便为插值提供更可靠的参考值。

图 1 展示了相邻 $Q (Q=4)$ 个子载波的信道估计过程, 扩展到 $Q > 4$ 情况下的原理与之相同。在不失一般性的情况下, 将相邻的 4 个子载波分别设置为 K_0 、 K_0+1 、 K_0+2 、 K_0+3 。假设最差的情况, 即 $\mathbf{W}_k = \mathbf{W}$, $\mathbf{F}_k = \mathbf{F}$, $\mathbf{X}_k = \sqrt{P} \mathbf{I}$, 其中, P 表示信号发射的功率, 则接收到的导频信号 \mathbf{Y}_k 为

$$\mathbf{Y}_k = \sqrt{P} \mathbf{W}^H \mathbf{H}_k \mathbf{F} + \tilde{\mathbf{N}}_k \quad (6)$$

将导频信号 \mathbf{Y}_k 送入数据预处理模块, 即

$$\mathbf{R}_k = \mathbf{Q} \mathbf{Y}_k \quad (7)$$

其中, \mathbf{Q} 为

$$\mathbf{Q} = (\mathbf{W} \mathbf{W}^H)^{-1} \mathbf{W} \quad (8)$$

综上可得

$$\mathbf{R}_k = \mathbf{Q} \mathbf{Y}_k = (\mathbf{W} \mathbf{W}^H)^{-1} \mathbf{W} (\sqrt{P} \mathbf{W}^H \mathbf{H}_k \mathbf{F} + \tilde{\mathbf{N}}_k) = \sqrt{P} \mathbf{H}_k \mathbf{F} + \mathbf{N}_k \quad (9)$$

2 基于 AAT 模型的信道估计

2.1 集中注意力机制网络

在时变信道的信道估计中, 具有多特征的信道数据, 可能存在部分特征与实际预测结果关联性小, 部分异常值对预测结果影响大, 以及信道特征提取不可持续影响预测结果等情况。如果直接使用时序卷积神经网络对信道数据进行处理, 效果可能会不理想。对于上述存在的问题, 在现有的 TCN 模型^[9]中融合集中注意力机制, 使之能够更加合理有效地提取数据特征, 在 AAT 模型中, 将信道矩阵视为带有噪声的图像信息, 因此考虑将图像处理中算法^[16]引入信道估计中。本文提出的 AAN 结构如图 2 所示。该注意力网络由注意力分支、非注意力分支以及一个可学习的动态注意力分配模块构成。与传统注意力机制不同的是, 本文构建一个注意力分配模块, 该模块能够为 2 个分支分配动态注意力权值, 能够自动丢弃部分不重要的注意力特征, 更加专注于有益的特征, 并且仅增加少量的额外参数就提升了注意力模型的容量。当输入为 x_n 时, AAN 的输出为

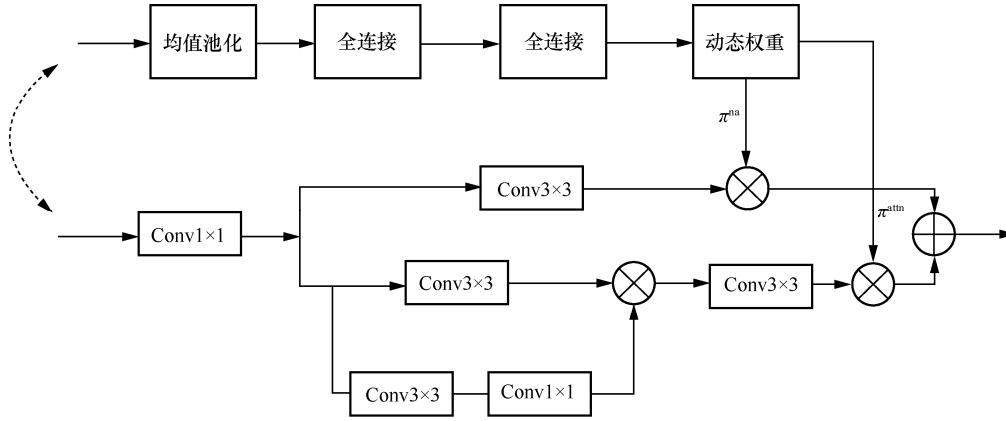


图 2 AAN 结构

$$x_{n+1} = f_{|x|}(\pi_n^{na} x_n^{na} + \pi_n^{attn} x_n^{attn}) \quad (10)$$

其中, $f_{|x|}(\cdot)$ 为卷积核大小为 1×1 的卷积, x_n^{na} 为非注意力分支的输出, x_n^{attn} 为注意力分支的输出, π_n^{na} 为非注意力分支的注意力权重, π_n^{attn} 为注意力分支的注意力权重。这 2 个权重是由网络根据输入数据的特征计算而得的, 并不是人工设置的。注意力权重的计算式为

$$\pi_n = f_{da}(x_n) \quad (11)$$

其中, $f_{da}(\cdot)$ 是注意力分配模块。首先, 使用全局均值池化来压缩输入 x_{n-1} ; 然后, 输入由 2 个全连接层组成的连接层中, 激活函数使用 ReLU。使用全局池化来增加注意力分配模块的感受野, 这使注意力分配模块能够从输入的整个图像(信道数据)获取特征。根据文献[13], 约束动态权重值可以更好地促进注意力分配模块的学习, 设置约束条件如下

$$\pi_n^{na} + \pi_n^{attn} = 1 \quad (12)$$

这里选用 Softmax 函数来生成注意力分支和非注意力分支的归一化权重, 式(1)约束动态权重和为 1 的方法可以简化 π_n 的学习。

2.2 AAT 模型

由文献[17]可知, 连续相干区间中的时变信道建模为高斯-马尔可夫分布, 表示为

$$\mathbf{H}_k[n] = \rho \mathbf{H}_k[n-1] + \sqrt{1-\rho^2} \boldsymbol{\Theta}[n], \quad n \in \mathbb{N}_+ \quad (13)$$

其中, $\mathbf{H}_k[n]$ 为 $\mathbf{H}_k(t)$ 的离散时间表示, $0 \leq \rho \leq 1$ 为时间相关系数, $\boldsymbol{\Theta}[n]$ 为一个用来说明每个条目单位差异的随机矩阵。式(13)表明从先前相干区间到当前相干区间的信道变化受到了一定的干扰。因此考

虑利用信道的时相关性提高信道估计的精度。信道估计流程如图 3 所示, 为简化说明, 考虑 $Q(Q=2)$ 个相邻子载波的情况, 当输入的信道通过系统模型时, 可以得到

$$\mathbf{R}_k[n] = \sqrt{P} \mathbf{H}_k[n] \mathbf{F} + \mathbf{N}_k[n] \quad (14)$$

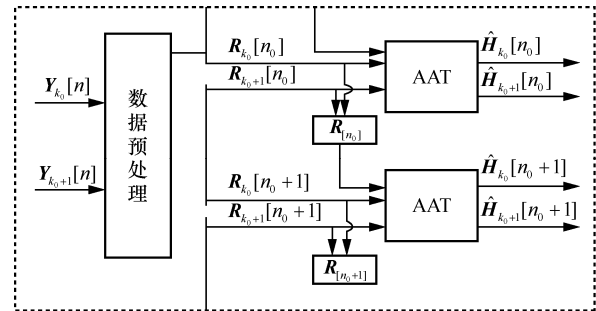


图 3 信道估计流程

AAT 模型结构如图 4 所示, 其中, Layer1 和 Layer2 分别表示输入的复信道矩阵的实部实数矩阵层和虚部实数矩阵层, 批归一化 (BN) 层用于缓解梯度消失问题。通过使用 AAT 模型对导频信号 \mathbf{Y}_k 进行进一步细化处理。采用第 n_0 和第 $(n_0 + 1)$ 这 2 个连续相干区间的情况来描述信道估计的过程。在第 n_0 个相干区间中, 暂时估计的信道矩阵 $\mathbf{R}_{k_0}[n_0]$ 和 $\mathbf{R}_{k_0+1}[n_0]$ 被输入 AAT 模型中。 $\mathbf{R}_{k_0}[n_0]$ 和 $\mathbf{R}_{k_0+1}[n_0]$ 的副本存储在存储器中以便用于下一个时间间隔中。在第 $(n_0 + 1)$ 个时间间隔中, AAT 模型接收临时的估计信道 $\mathbf{R}_{k_0}[n_0 + 1]$ 和 $\mathbf{R}_{k_0+1}[n_0 + 1]$, 并获取先前的存储副本 $\mathbf{R}_{k_0}[n_0]$ 进行联合处理, 获得想要的估计信道 $\hat{\mathbf{H}}$ 。此时将存储器中的 $\mathbf{R}_{k_0}[n_0]$ 更新为 $\mathbf{R}_{k_0}[n_0 + 1]$, 即 $\mathbf{R}_{k_0}[n_0 + 1]$ 和 $\mathbf{R}_{k_0+1}[n_0 + 1]$ 。本文设计该模型的目的是学习时变信道的一般特征, 而不是 2 个连续相干

区间中信道的特有特征。因此在每一个相干间隔中都使用相同结构的 AAT 模型。

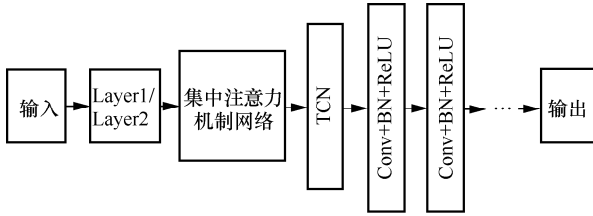


图 4 AAT 模型结构

在使用 AAT 模型对信道进行估计之前，要对 AAT 模型进行离线训练。首先，将带有噪声的临时估计信道矩阵 \mathbf{R}_{k_0} 作为输入，输入先前设计好的集中注意力机制网络中，该网络通过注意力分支以及非注意力分支能够自动丢弃输入信道中的非重要特征和噪声，而重点学习时变信道中的固有特征。集中注意力机制网络将输入的信道矩阵作为图像来进行处理，对每一层特征图进行采样直接传输到下一层中。在集中注意力机制网络后面，本文设计了 TCN 对收到的信道图像进行特征提取。对输入的信道矩阵使用 64 个 $3 \times 3 \times 4$ 卷积滤波器以及 ReLU 激活函数进行处理，生成 64 个 16×32 的实值矩阵。最后设置一个输出层，通过使用 Tanh 函数作为激活函数将输出映射到区间 $[-1, 1]$ 。将相应的实部与虚部放大组合之后，即可得到想要的估计信道矩阵 $\hat{\mathbf{H}}$ 。

3 实验及结果分析

3.1 数据集与实验设置

3GPP 为 5G 系统定义了一种新的三维无线电信道模型规范。该模型支持 $0.5 \sim 100$ GHz 的载波频率，系统的带宽高达 2 GHz，并且充分研究了信道的随机特性。因此本文选用 3GPP TR38.901 Release15 信道模型来生成信道数据^[18]。该模型的信道建模流程如图 5 所示，包括生成基本参数、生成小尺度参数和生成信道系数的过程。

本文所采用的城市微小区 (UMi) 场景参数设置如表 1 所示。系统采用簇延时线性模型，载波频率 $f_c = 28$ GHz，采样率 $f_s = 100$ MHz，主路径数 $L = 1$ ，子载波数 $K = 64$ 。发射天线数和接收天线数分别为 $N_T = 32$ ， $N_R = 16$ ，发射机和接收机的射频链数为 $N_T^{RF} = N_R^{RF} = 2$ 。同时将 \mathbf{F} 和 \mathbf{W} 设置为 $N_T \times N_T$ 的离散傅里叶变换 (DFT) 矩阵的第 M_T 列和 $N_R \times N_R$ 的 DFT 矩阵的第 M_R 列。

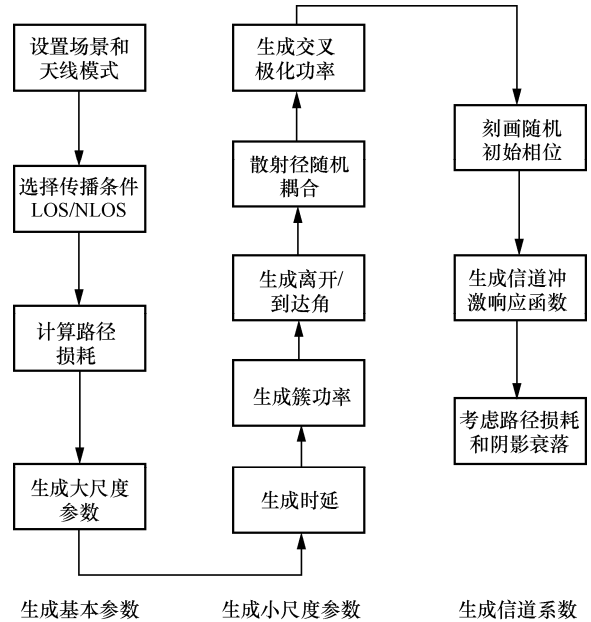


图 5 信道建模流程

表 1 UMi 场景参数设置

名称	值
主路径	1, 2, 3, 4
载波频率/GHz	28
采样率/MHz	100
子载波数	64
发射天线数	32
接收天线数	16
最大多普勒频移	0
天线阵列	ULA

AAT 模型详细参数如表 2 所示。训练集、测试集、验证集分别包含 90 000、10 000、10 000 个样本。epoch 设置为 800，缩放常数 $c = 2$ ，学习率 $r = 10^{-5}$ 。

表 2 AAT 模型详细参数

参数	值
训练集样本数	90 000
验证集样本数	10 000
测试集样本数	10 000
轮数	800
批量大小	128
缩放常数	2
学习率	10^{-5}
卷积层激活函数	ReLU
输出层激活函数	Tanh

3.2 评估标准

为了客观评估本文所提出的 AAT 模型和其他算法的性能，采用归一化均方误差（NMSE，normalized mean square error）作为评价指标进行分析。NMSE 表达式为

$$NMSE = \mathbb{E}_H \left\{ \frac{\| \mathbf{H} - \widehat{\mathbf{H}} \|_F^2}{\| \mathbf{H} \|_F^2} \right\} \quad (15)$$

其中， \mathbf{H} 和 $\widehat{\mathbf{H}}$ 分别表示真实信道和估计信道。

3.3 仿真分析

3.3.1 AAT 模型在不同卷积层下的性能

本文在 AAT 模型结构的设计前期，对卷积层数的设置进行了研究，分别对 AAT、AAT+1 层卷积和 AAT+5 层卷积场景下的 NMSE 和复杂度进行仿真实验，实验结果如图 6 及表 3 所示。结果表明，进一步增加卷积层数不能带来显著的性能增益，反而会导致网络训练和测试的复杂度大幅提升。

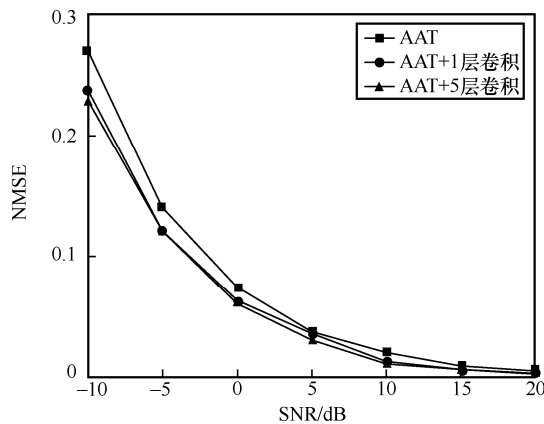


图 6 不同卷积层数下 AAT 模型的 NMSE

表 3 不同卷积层数下的 AAT 模型的复杂度

网络模型	参数量	浮点运算量
AAT	0.55×10^6	63 523 364
AAT+1Conv	0.58×10^6	71 337 702
AAT+5Conv	0.7×10^6	81 664 533

3.3.2 神经网络模型的性能分析

考虑到信道的时延相关性，使用 LSTM 网络能够充分利用信道的这一特性。本节通过比较 TCN 以及 LSTM 网络在信道估计中的精度以及复杂度，来解释本文为何选择 TCN 进行改进而不是 LSTM 网络。不同网络在不同信噪比下的性能如图 7 所示。在相同的子载波以及相干间隔下，TCN、LSTM 以及 Bi-LSTM 网络在不同信噪比下的性能几乎相同。

当不同网络的性能非常接近时，复杂度就成为选择网络的重要指标。

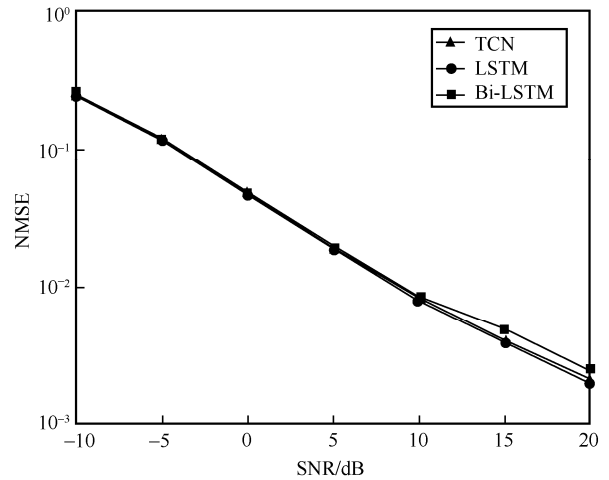


图 7 不同网络在不同信噪比下的性能

神经网络模型的时间复杂度和计算复杂度分别如表 4 和表 5 所示。从表 4 可以看出，LSTM 和 Bi-LSTM 网络的训练时间和测试时间都远远超过 TCN。从表 5 可以看出，TCN 相较于 LSTM 和 Bi-LSTM 网络具有更低的参数量以及计算复杂度，这表明 TCN 网络具有更高的训练效率。

表 4 神经网络模型的时间复杂度

网络模型	训练时间/s	测试时间/s
TCN	21	1.5×10^{-4}
LSTM	33	3.5×10^{-4}
Bi-LSTM	68	7.1×10^{-4}

表 5 神经网络模型的计算复杂度

网络模型	参数量	浮点运算量
TCN	0.5×10^6	154 435 805
LSTM	0.7×10^6	201 447 752
Bi-LSTM	1.4×10^6	381 226 002

综上所述，LSTM 以及 Bi-LSTM 网络在信道估计性能上较 TCN 并没有性能优势，但时间复杂度和计算复杂度都要高很多。由于 TCN 能够有效提取信道的时延相关性信息，并且复杂度相对较低，因此本文选用 TCN 作为基础网络。

3.3.3 基于 AAN 的 5 种时序卷积神经网络模型的性能比较

本节在上述常见的深度学习网络中嵌入集中注意力网络，并与本文所提网络进行性能以及复杂度对比，对比结果如表 6、表 7 和图 8 所示。

表 6 基于 AAN 的 5 种时序卷积神经网络的时间复杂度

网络模型	训练时间/s	测试时间/s
AAT	29	3×10^{-4}
LSTM-AAN	60	7×10^{-4}
Bi-LSTM-AAN	115	13×10^{-4}
RNN-AAN	41	4×10^{-4}
GRU-AAN	55	5.4×10^{-4}

表 7 基于 AAN 的 5 种时序卷积神经网络的计算复杂度

网络模型	参数量	浮点运算量
AAT	0.55×10^6	63 523 364
LSTM-AAN	0.75×10^6	102 445 677
Bi-LSTM-AAN	1.45×10^6	204 337 811
RNN-AAN	0.67×10^6	872 335 776
GRU-AAN	0.86×10^6	112 448 755

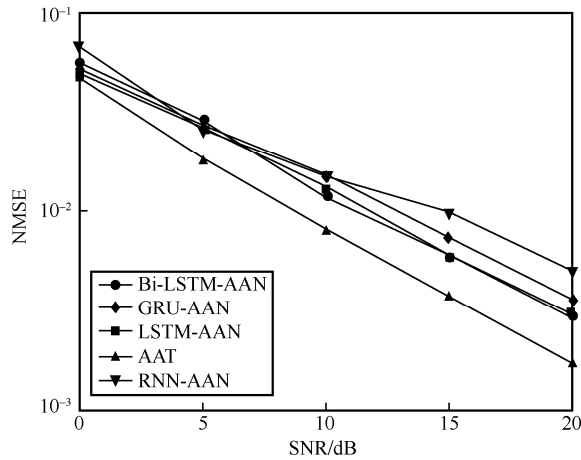


图 8 基于 AAN 的 5 种时序卷积神经网络模型在不同信噪比下的性能

从表 6、表 7 和图 8 可以看出，AAT 模型相较于其他深度学习网络在提升性能的同时，依然保持着较低的复杂度。

3.3.4 不同主路径数下的网络鲁棒性分析

在真实的场景中，主路径数 L 可能会发生变化。因此本文测试了基于 AAT 模型的方法对不同主路径数的鲁棒性，如图 9 所示。 $L=3$ 是在训练阶段测得的数据， $L=1,2,4$ 是在测试阶段测得的数据。测试阶段 $L=3$ 的情况与训练阶段相匹配，因此选其为比较的基准。此外，选用传统的 LS 信道估计方法作为比较的上限。仿真结果表明，基于 $L=1$ 的 AAT 模型方法比基线性能更好，这是因为信道结构较简单，有利于卷积神经网络去噪，当 L 增加到 2 时性能相比基线略有下降，虽然当 L 增加到 4 的时候，信道估计的性能出现了明显的下降，

但当 $\text{SNR} < 20 \text{ dB}$ 时，该方法仍优于 Bi-LSTM-AAN、LSTM-AAN。

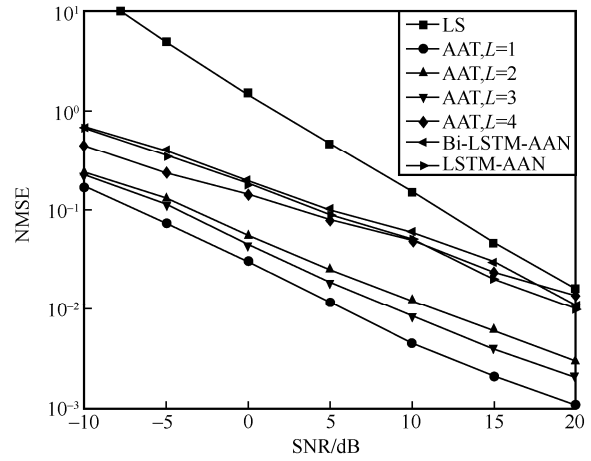


图 9 AAT 模型在不同主路径数下的 NMSE

真实场景中，基于 TCN 模型的方法在不同主路径数下的 NMSE 如图 10 所示。对比图 9 和图 10，在主路径数不匹配的情况下，融合了集中注意力机制的 AAT 模型仍有利于提高信道估计的精度。

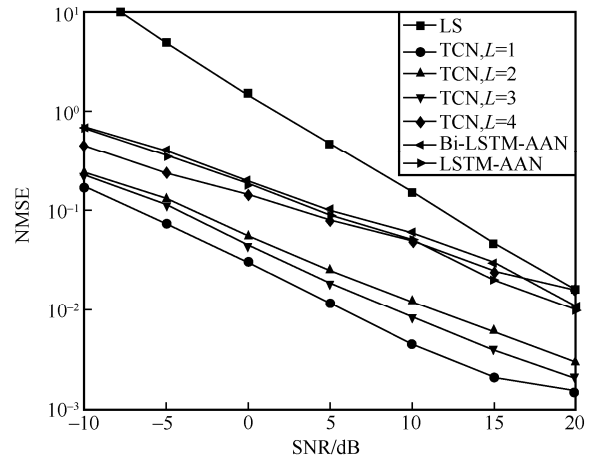


图 10 TCN 模型在不同主路径数下的 NMSE

3.3.5 不同场景下的网络性能分析

当最大多普勒频移 $f_d = 1400 \text{ Hz}$ 时，在 UMi 街道 NLOS 场景下进行训练，并测试不同模型的信道估计性能。基于 CNN、TCN、AAT 和 MMSE 信道估计方法的 NMSE 与 SNR 的关系如图 11 所示。MMSE、CNN、TCN 和 AAT 都是利用先前相干间隔的信道信息来联合估计 2 个子载波上的当前信道。比较 TCN 和 AAT 可以发现，AAT 模型中的联合注意力机制更好地提取了相邻子载波和

相干间隔的附加信道信息，显著提高了信道估计的精度。在相同的信道相关信息下，基于 AAT 以及 TCN 的方法在性能以及复杂度上都显著优于传统的 MMSE，代价仅是需要先进行离线训练。

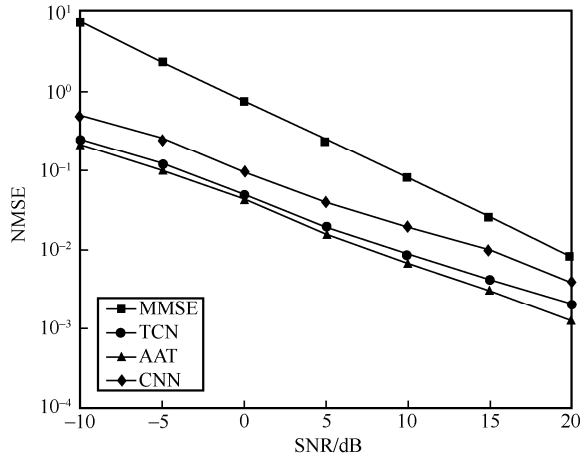


图 11 在 UMi 街道 NLOS 场景下不同方法的 NMSE 与 SNR 的关系

当最大多普勒传播 $f_d = 1\ 800$ Hz 时，在城市宏小区 (UMa) 街道 NLOS 场景下不同方法的 NMSE 与 SNR 的关系如图 12 所示。与图 11 相比，图 12 中 MMSE 的性能发生了严重的衰退，这是因为在数据预处理阶段计算 LS 估计信道时，协方差矩阵不匹配。由于 TCN、CNN、AAT 模型学习了信道固有的特征结构，因此对显著不同的测试场景依然具有优秀的鲁棒性，并且在这种情况下，融入了集中注意力机制的 AAT 模型依旧比 TCN 具有更优秀的性能。

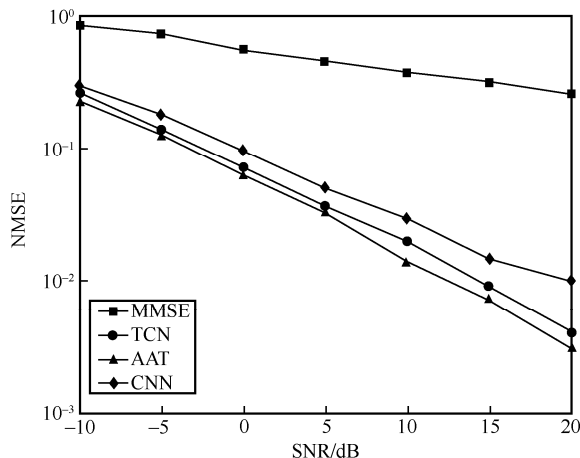


图 12 在 UMa 街道 NLOS 场景下不同方法的 NMSE 与 SNR 的关系

4 结束语

考虑到 MIMO 信道的时间相关性以及注意力

集中机制强大的特征提取能力，本文提出了一种融合时序卷积神经网络和集中注意力机制网络的 AAT 信道估计模型。该模型利用时序卷积神经网络对信道数据前后向的感知能力，注意力集中机制根据不同的输入特征对注意力进行合理的动态分配，充分挖掘了数据集中所包含的重要信息，使该模型具有比 TCN 模型更好的特征提取能力。实验数据表明，AAT 模型相较于 TCN、CNN 能够更好地捕获信道数据中的深层特征，并且在测试场景发生改变时，仍然能够保持很好的性能。这为降低毫米波多输入多输出系统的复杂性和成本提供了一种可行的方案。

参考文献:

- [1] ZHAO J H, NI S J, YANG L H, et al. Multiband cooperation for 5G HetNets: a promising network paradigm[J]. IEEE Vehicular Technology Magazine, 2019, 14(4): 85-93.
- [2] ABDALLAH A, CELIK A, MANSOUR M M, et al. Deep learning-based frequency-selective channel estimation for hybrid mmWave MIMO systems[J]. IEEE Transactions on Wireless Communications, 2022, 21(6): 3804-3821.
- [3] LI X F, ALKHATEEB A. Deep learning for direct hybrid precoding in millimeter wave massive MIMO systems[C]//Proceedings of the 2019 53rd Asilomar Conference on Signals, Systems, and Computers. Piscataway: IEEE Press, 2019: 800-805.
- [4] DONG P H, ZHANG H, LI G Y. Framework on deep learning-based joint hybrid processing for mmWave massive MIMO systems[J]. IEEE Access, 2020, 8: 106023-106035.
- [5] QIAO Y T, YU S Y, SU P C, et al. Research on an iterative algorithm of LS channel estimation in MIMO OFDM systems[C]//Proceedings of the IEEE 6th Circuits and Systems Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communication. Piscataway: IEEE Press, 2004: 729-732.
- [6] RAO X B, LAU V K N. Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems[J]. IEEE Transactions on Signal Processing, 2014, 62(12): 3261-3271.
- [7] FAN D, GAO F F, LIU Y W, et al. Angle domain channel estimation in hybrid millimeter wave massive MIMO systems[J]. IEEE Transactions on Wireless Communications, 2018, 17(12): 8165-8179.
- [8] DUONG N S, NGUYEN Q T, NGO K H, et al. Sparse Bayesian learning with atom refinement for mmWave MIMO channel estimation[C]//Proceedings of the 2023 IEEE Statistical Signal Processing Workshop (SSP). Piscataway: IEEE Press, 2023: 155-159.
- [9] DONG P H, ZHANG H, LI G Y, et al. Deep CNN-based channel estimation for mmWave massive MIMO systems[J]. IEEE Journal of Selected Topics in Signal Processing, 2019, 13(5): 989-1000.
- [10] 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.
- [11] DONG Y D, WANG H X, YAO Y D. Channel estimation for one-bit

multiuser massive MIMO using conditional GAN[J]. IEEE Communications Letters, 2021, 25(3): 854-858.

- [12] RUAN C Y, ZHANG Z C, JIANG H, et al. Simplified learned approximate message passing network for beamspace channel estimation in mmWave massive MIMO systems[J]. IEEE Transactions on Wireless Communications, 2023, PP(99): 1.
- [13] ZHAO J H, WU Y, ZHANG Q M, et al. Two-stage channel estimation for mmWave massive MIMO systems based on ResNet-UNet[J]. IEEE Systems Journal, 2023, 17(3): 4291-4300.
- [14] GAO Z, HU C, DAI L L, et al. Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective fading channels[J]. IEEE Communications Letters, 2016, 20(6): 1259-1262.
- [15] QIN Z J, YE H, LI G Y, et al. Deep learning in physical layer communications[J]. IEEE Wireless Communications, 2019, 26(2): 93-99.
- [16] CHEN H, GU J, ZHANG J. Attention in attention network for image super-resolution[J]. arXiv Preprint, arXiv: 2104.09497, 2021.
- [17] VA V, CHOI J, HEATH R W. The impact of beamwidth on temporal channel variation in vehicular channels and its implications[J]. IEEE Transactions on Vehicular Technology, 2017, 66(6): 5014-5029.
- [18] 邢冰倩, 韦再雪. 基于 3GPP TR38.901 信道模型的时变信道建模与仿真[J]. 北京邮电大学学报, 2021, 44(1): 45-51.
- XING B Q, WEI Z X. Time-varying channel modeling and simulation based on 3GPP TR38.901 channel model[J]. Journal of Beijing University of Posts and Telecommunications, 2021, 44(1): 45-51.

[作者简介]



于舒娟（1967-），女，江苏南京人，南京邮电大学教授，主要研究方向为自适应信号处理、深度学习和智能大数据处理。



刘荣（1999-），男，江苏南通人，南京邮电大学硕士生，主要研究方向为深度学习与信号处理。



张昀（1975-），女，江苏南京人，博士，南京邮电大学副教授，主要研究方向为智能化算法与通信信号处理。



谢娜（1980-），女，江苏南京人，博士，南京邮电大学副教授，主要研究方向为智能化算法与通信信号处理。



黄丽亚（1972-），女，湖南绥宁人，博士，南京邮电大学教授，主要研究方向为通信网络的 QoS 性能研究自适应信号处理、深度学习和智能大数据处理。