Implementation of Adapter-Based Transfer Learning for Channel Estimation Using a Pretrained Transformer in Multi-Channel Environments

Joo Won Lee, Kae Won Choi*

Sungkyunkwan Univ., *Sungkyunkwan Univ.

joowonoil@skku.edu, *kaewonchoi@skku.edu

다중 채널 환경에서 사전 학습된 Transformer 기반 채널 예측 모델의 어댑터 전이학습 구현 및 성능 평가에 대한 연구

이주원, 최계원*
성균관대학교, *성균관대학교

Abstract

This paper presents an adapter-based transfer learning approach for channel estimation using a Transformer pretrained on InF-LoS, InF-NLoS, and RMa-LoS environments. Only lightweight adapters are fine-tuned for InF and RMa-LoS while core layers are frozen. Results show improved accuracy with environment-specific tuning and demonstrate both the efficiency and generalization ability of the model across diverse environments.

Ⅰ. Introduction

Accurate estimation of channel state information (CSI) is essential for enabling efficient communication in next-generation wireless systems [1]. Recently, deep learning approaches, particularly Transformer-based models, have demonstrated promising results in capturing complex temporal and spatial characteristics of wireless channels [2]. However, most existing models are trained and evaluated in single-channel environments, limiting their generalization performance when applied to new conditions. This motivates the need for studying transferability across diverse channel scenarios [3].

In this paper, we pretrain a Transformer-based CSI estimation model using data from InF-LoS, InF-NLoS, and RMa-LoS environments. We then fine-tune the model for specific environments—InF and RMa-LoS—and evaluate its inference performance in both. The results provide insights into how transfer learning impacts the generalization capability of CSI predictors in practical multi-environment deployments.

Ⅱ. Method

A. Dataset Construction

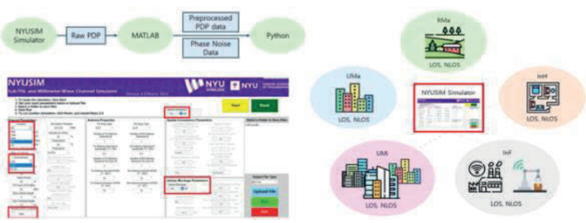

Fig 1. PDP Dataset generation from NYUSIM

NYUSIM is used to simulate frequency-selective wireless channels under three environments: InF-LoS, InF-NLoS, and RMa-LoS. Each scenario yields 50,000 PDPs at 28 GHz, aligned to zero delay at the first path. These are converted into CIRs and then transformed into frequency-domain CSI. Input reference signals follow the 5G NR standard, with DMRS pilots inserted every 7 subcarriers. Each input tensor includes both real and imaginary parts. To ensure unbiased training, the dataset is divided into four disjoint subsets: pretraining, fine-tuning, validation, and inference. A data loader is used to supply input-output pairs during model training.

B. Pretraining Configuration

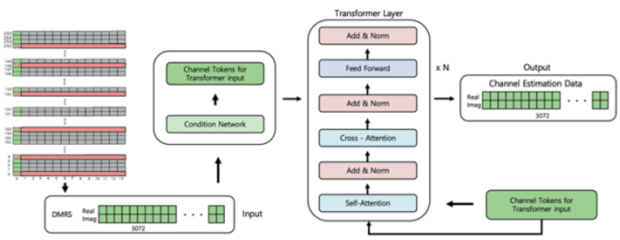

Fig 2. Transformer-based NN Model Architecture

To pretrain a generalized channel estimation model, we employ a Transformer encoder architecture optimized for frequency-selective wireless channels. The model consists of four encoder layers, each equipped with eight multi-head attention blocks and a hidden dimension of 128. Each encoder also includes a position-wise feed-forward network with a width of 1024 to capture local representations. Positional encoding is added to preserve the frequency-domain structure of the input signals, and dropout with a rate of 0.1 is applied to prevent overfitting. The model is trained using the Adam optimizer over 3 million iterations. A fixed learning rate of 1e-5 is used, along with a weight decay of 1e-6 to regularize the training. Gradient clipping with a maximum norm of 1.0 is applied to ensure stable convergence. These settings are chosen to balance model expressiveness with training stability, enabling the model to effectively learn from large-scale simulated datasets across multiple environments.

C. Transfer Learning with Adapters

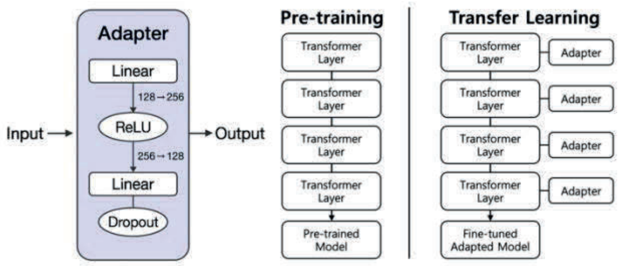

Fig 3. Comparison of pre-trained and adapter-augmented Transformer architectures for environment-specific transfer learning.

To adapt a pretrained Transformer model to specific channel environments, we employ an adapter-based transfer learning strategy. Instead of fine-tuning the entire model, we freeze all four encoder layers and insert lightweight adapter modules at the end of each layer. This approach preserves the generalized representations learned during pretraining while allowing efficient environment-specific adaptation. Each adapter module is a two-layer feed-forward network with a bottleneck architecture: the output of each Transformer layer is first compressed to 256 dimensions, activated by ReLU, and then projected back to the original dimension. A dropout rate of 0.1 is applied within the adapter to improve generalization performance. We perform fine-tuning separately for the InF and RMa-LoS environments. During this process, only the adapter parameters are updated; the rest of the model remains fixed. Training is conducted for 100,000 iterations using the Adam optimizer with a learning rate of 1e-4, weight decay of 1e-6, and gradient clipping with a max norm of 1.0. Early stopping is applied with a patience of 500 iterations and a minimum performance delta of 0.0001 to prevent overfitting. All training is performed on a CUDA-enabled GPU, and experiments are tracked using the Weights & Biases (W&B) platform to ensure reproducibility and monitoring throughout the training process.

Ⅲ. Conclusion

In this work, we implemented and evaluated an adapter-based transfer learning approach for channel estimation using a Transformer model pretrained on multiple wireless channel environments, including InF-LoS, InF-NLoS, and RMa-LoS. By freezing the core Transformer layers and fine-tuning only lightweight adapters, we efficiently adapted the model to specific environments such as InF and RMa-LoS.

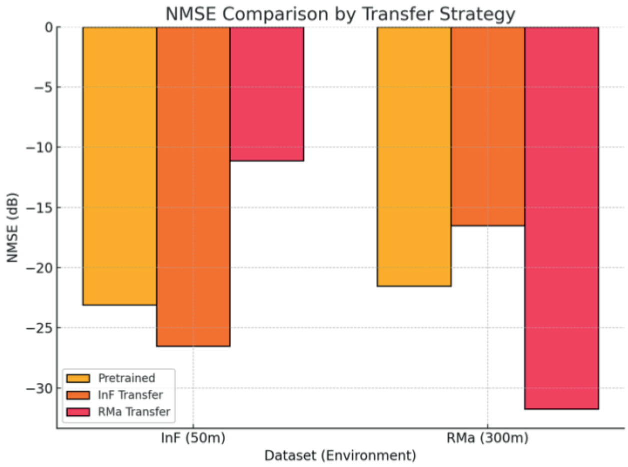

Fig 4. NMSE(dB) Comparison of pre-trained and adapter-based Transformer models on InF(50m) and RMa(300m) environments.

Experimental results (Fig 4.), obtained in InF and RMa settings at 50 meters and 300 meters respectively, demonstrate that the adapter-based fine-tuning significantly improves estimation accuracy in the target domain. Moreover, the method effectively transfers knowledge from the generalized pretrained model to each specific environment, achieving a favorable balance between adaptation and generalization. These findings suggest that the proposed framework is both effective and scalable for environment-aware channel estimation. For future work, we plan to investigate the impact of adapter structure variations—such as changes in bottleneck dimension or insertion position—as well as the effect of increasing the amount of pretraining data on model generalization and downstream performance.

ACKNOWLEDGMENT

This work was supported by the BK21 FOUR Project.

REFERENCES

[1] H. Huang, C. Qi, C. Jiang, and G. Y. Li, “Deep learning for physical-layer 5G wireless techniques: Opportunities, challenges and solutions,” IEEE Wireless Communications, vol. 27, no. 1, pp. 138-145, Feb. 2020.

[2] C. Huangxun, Y. Wang, L. Chen, Z. Wang, and H. Liu, “Cross-shaped Separated Spatial-Temporal UNet Transformer for Accurate Channel Prediction,” Proc. IEEE INFOCOM, Apr. 2024.

[3] Y. Yang, F. Gao, Z. Zhong, B. Ai, and A. Alkhateeb, “Deep Transfer Learning Based Downlink Channel Prediction for FDD Massive MIMO Systems,” arXiv preprint arXiv:1912.12265, Dec. 2019.