

Peer Review

Review of: "Large Models Enabled Ubiquitous Wireless Sensing"

Fei Zhang¹

1. Huawei Technologies, Shenzhen, China

The authors tried to make cross-model comparisons for spatial channel prediction using ray-tracing channel data generated by commercial software. However, the work lacks necessary insight. Based on the presentation, it is difficult to see the contribution.

Problem definition: based on the description, it is not clear what the CSI to be predicted is. For example, CSIs are denoted in three different ways in equations 1, 2, and 3; it is not clear what the connections between them are, namely $H(t, \tau)$, h_k^H , H_d . How is time S defined? Is it a slot number or an RB number? What are the exact values of L and P in the evaluation? What does it exactly mean by spatial CSI? How is it defined analytically? Later, for the CSI data pre-processing section, the complex channel matrix is defined as H_f . What is f ?

Model training: what is fed into the model? And how? When using the pre-trained GPT-2, which layers are frozen? Which are trainable? By only saying "most layers" is not clear. How is training for the other NNs done? Are they pre-trained? How are the weights initialized? How much data is used for training each of them?

Why is the loss function defined as it is in equation 4? Some reasoning is needed.

Pre-processing of WinProp CSI: firstly, the text in the upper branch of Figure 8 is barely readable. What is the motivation to incorporate hand-crafted features into the NN instead of working with the original CSI?

Results: it is difficult to arrive at a comparison between all the NNs considered since the results are shown in separate figures. The VAE shows significantly better results than the others. The transformer, GPT-2, and MLP are surprisingly worse. However, no analysis has been provided: why are these NNs considered for evaluation? Why do they deliver performance like shown? When

evaluating a pre-trained model like GPT-2, fine-tuning is required for deterministic channel data. It is not clear how this is done.

In general, it is an interesting direction to explore NN performance based on deterministic channel data in order to understand robustness across different data generation methods when training/evaluating NNs. The authors are encouraged to provide clear analysis linking different NN performance to deterministic channel properties.

Declarations

Potential competing interests: No potential competing interests to declare.