IMPROVING CDL CHANNEL MODEL FIDELITY THROUGH EXTRACTING CHANNEL CHARACTERISTICS FROM LIMITED CSI LOGS USING MACHINE LEARNING

Benjamin Earle

benearle@cmail.carleton.ca

SUMMARY

Over The Air (OTA) testing for wireless systems is expensive and time consuming. To save Research and Development (R&D) cost, testing is done in simulation prior to conducting OTA tests. Cluster Delay Line (CDL) channel models are widely used in state-of-the-art wireless simulations; however, their configuration is often limited to one of the five pre-set models. The pre-set models are intended for benchmarking, they have insufficient coverage for pre-deployment testing. Improving CDL model fidelity is crucial to reduce the amount of OTA testing required during R&D. This research demonstrates a methodology to recreate real channels in simulation, in a CDL channel model. This is done through creating a custom CDL profile from Channel State Information (CSI) logs with Machine Learning (ML). There is no real ground truth data for this problem, so supervised ML models are trained using synthetic data generated by a custom wireless simulator. Once trained the ML models will recreate real channels in simulation using CSI collected during OTA tests.

Keywords: Wireless Channel Reconstruction, Cluster Delay Line, and Machine Learning.

ACKNOWLEDGEMENTS

I would like to acknowledge my thesis co-supervisors Dr. X and Dr. X for guiding my research. Additionally, I would like to acknowledge Ericsson Canada for their partnership and funding this research. Namely, I would like to thank Dr. X and Dr. X from Ericsson for providing wireless expertise and guiding the research to ensure it is focused on practical problems.

Proc. of the 2024 Annual Simulation Conference (ANNSIM'24), May 20-23, 2024, American University, DC, USA
P.J. Giabbanelli, I. David, C. Ruiz-Martin, B. Oakes and R. Cárdenas, eds.

PHD COLLOQUIUM ABSTRACT

Wireless communication Research and Development (R&D) is a massive industry that needs constant innovation to meet the ever rising demand. System testing is a critical component of the wireless R&D life cycle that must be streamlined to improve profits and time to market. Over-The-Air (OTA) testing of wireless networks is time consuming and expensive. It requires many people, completely configured equipment, and large, dynamic deployment areas. As a result of these challenges, wireless R&D relies heavily on simulated testing to validate systems prior to investing in OTA testing [1]. The simulated model fidelity is important to ensure that the systems are aptly vetted. The higher the simulation accuracy, the fewer the discrepancy between simulated and actual performance. Ultimately, through improving wireless simulations the amount of time spent conducting OTA tests will be reduced and as a result the time and cost of R&D. This PhD research proposes a methodology to improve wireless simulation accuracy through recreating real channels experienced in deployed systems. Existing works have studied analytical techniques to estimate simulation model parameters using ray tracing [2] [3]. These analytical solutions require a 3D map and complete knowledge of the physical environment, which is not widely available and is subject to change over time. The Machine Learning (ML) approach taken in this thesis will extract the simulation parameters from data that is easily collected from any deployed wireless network.

This work is a part of a collaborative effort with Ericsson Canada, with the objective of improving wireless channel model fidelity by using Machine Learning (ML) to tune the simulated model parameters. The Cluster Delay Line (CDL) channel model is a state-of-the-art model developed by 3GPP for 5G and beyond (5GB) networks [4]. CDL models offer realistic multi-path representation through modeling each path between the transmitter and receiver as a cluster of rays. They are adaptable to many propagation environments and frequency ranges. The models are a hybrid between stochastic and geometric channel models; they are commonly used in industry for MIMO link level system simulation due to their balance between efficiency and accuracy. However, configuring and generating the parameters for these models is non-trivial, and as such, many studies test their systems using only the pre-set configurations. The five pre-sets do not cover all potential wireless scenarios, and hence, should only be used as a benchmark. This PhD thesis presents a method to create physically accurate CDL profiles using ML on data collected from OTA tests.

In practice, the BS optimizes its communication with the UE through channel sounding [5]. This process is a part of the regular system operation and results in a set of Channel State Information (CSI) metrics that are used by the Base Station (BS) to optimize the communication link with the User Equiment (UE). It does not require much extra effort to store the CSI received at the BS, which makes it an ideal set of data to use when analyzing OTA channels. A subset of commonly available CSI was chosen as the input for our channel reconstruction ML models. The channel will be reconstructed as a CDL channel model through their several configurable parameters [4]. The channel parameters are split into three categories: System Configuration Parameters (SCP), Large Scale Parameters (LSP), and SSP. The aforementioned five pre-sets are example CDL model configurations and they offer a complete set of SCP, LSP, and SSP which model a specific environment. 3GPP also define a custom CDL profile generation algorithm to generate a new set of random parameters. The algorithm sequentially generates the SCP, LSP, and SSP; where in, each parameter is dependant the previously generated ones. The SCP consist of generic wireless system configurations, along with the presence of a line of sight cluster. There are seven LSP: shadow fading, K-factor, delay spread, and azimuth and zenith angle spread for arrival and departure angles. These traits will then influence the generation of the SSP, which are the individual cluster delays, powers, azimuth and zenith angles of arrival and departure, and the cross power ratios. The algorithm then uses the SSP to determine the channel matrix, which dictates the effect that will be applied on the signal within the simulation.

A given channel's characteristics and behaviour are not easily observable; therefore, synthetic training dataset is required to train the ML models. This dataset is generated using a simulation tool that spans the space of potential CDL profiles. The profiles are generated by adhering to a custom profile generation sequence defined by 3GPP [4]. The tool then simulates each of the CDL profiles to generate labeled data for the supervised learning studies. This dataset contains the CDL parameters (all of the SCP, LSP, and SSP) that were simulated along with the resulting CSI recorded over a series of channel sounding sequences. The real CSI calculation is not perfect due to random channel noise impacting the transmission. As such, the synthetic data includes a real and practical CSI values. Predicting the channel characteristics from noisy CSI readings proves to be more difficult, so we proposed a study using de-noising neural networks to improve the real CSI accuracy.

The machine learning models are trained using the synthetic data to estimate the channel characteristics present for each of the CSI readings. The models used in this research are primarily meta-learning combinations of deep neural networks and random forest regressors. Most SCP are known to the system, as they are static configuration parameters; with the exception of presence of a line of sight cluster, which can be accurately predicted from the CSI. To date the research has had mixed levels of success while estimating the LSP and SSP. Certain parameters can be estimated accurately, such as the direction of departure of dominant clusters from the BS. Other parameters have proven challenging and research is on going to improve the estimation results. There are some parameters where it is clear that the data present is insufficient for them to be accurately predicted. In these cases, the parameter will be generated from the random distribution recommended by 3GPP given the value of the other parameters.

Once the ML models are trained and validated on the synthetic data they will be ready to reconstruct real channels. The limited CSI chosen for their input is calculated constantly in real systems and can be collected easily. The ML models will predict the CDL channel profiles at both the LSP and SSP levels. The LSP prediction will allow for the creation of a class of channels, that could have different SSP. Where as the SSP prediction will create a specific channel experienced, minus the randomized background noise. These reconstructed channel profiles can then be simulated in CDL channel models, to recreate the experience environment in a simulation. This will result in improved CDL model fidelity by easily generating realistic CDL profiles. The increase in accuracy and availability of simulation environments will hopefully allow for more thorough testing and ultimately help reduce the amount of OTA testing required in wireless R&D.

REFERENCES

[1] M. S. Obaidat and D. B. Green, *Simulation of Wireless Networks*. Boston, MA: Springer US, 2003, pp. 115–153. [Online]. Available: https://doi.org/10.1007/978-1-4419-9218-5_6

[2] J. Yao, H. Ren, and Q. Liu, "Massive mimo channel modeling using map-based ray tracing method," in *2017 7th IEEE International Symposium on Microwave, Antenna, Propagation, and EMC Technologies (MAPE)*, 2017, pp. 1–5.

[3] R. Gentile, D. Garcia-Alis, and H. Chen, "Algorithms to antenna: Clustered-delay-line channel model with ray tracing," https://www.mwrf.com/technologies/embedded/systems/article/21155377/mathworks-algorithms-to-antenna-clustered-delay-line-channel-model-with-ray-tracing.

[4] T. 3rd Generation Partnership Project, "Study on channel model for frequencies from 0.5 to 100 ghz tr 38.901," 2020.

[5] E. Dahlman, S. Parkval, and J. Skold, *5G NR: The Next Generation Wireless Access Technology*. Academic Press, 2018.