IMPLICIT CHANNEL LEARNING FOR MACHINE LEARNING APPLICATIONS IN 6G WIRELESS NETWORKS

Ahmet M. Elbir\textsuperscript{1}, Wei Shi\textsuperscript{2}, Kumar Vijay Mishra\textsuperscript{1,3}, Anastasios K. Papazafeiropoulos\textsuperscript{1,4} and Symeon Chatzinotas\textsuperscript{1}

\textsuperscript{1}Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, Luxembourg
\textsuperscript{2}School of Information Technology, Carleton University, Ottawa, Canada
\textsuperscript{3}United States DEVCOM Army Research Laboratory, Adelphi, USA
\textsuperscript{4}University of Hertfordshire, Hatfield, U. K.

ABSTRACT

With the deployment of the fifth generation (5G) wireless systems gathering momentum across the world, possible technologies for 6G are under active research discussions. In particular, the role of machine learning (ML) in 6G is expected to enhance and aid emerging applications such as virtual and augmented reality, vehicular autonomy computer vision and internet of everything. This will result in large segments of wireless data traffic comprising image, video and speech. The ML algorithms process these for classification/estimation/through the learning models located on cloud servers. This requires wireless transmission of data from edge devices to the cloud server. Channel estimation, handled separately from recognition step, is critical for accurate learning performance. Toward combining the learning for both channel and the ML data, we introduce implicit channel learning to perform the ML tasks without estimating the wireless channel. Here, the ML models are trained with channel-corrupted datasets in place of nominal data. Without channel estimation, the proposed approach exhibits approximately 60% improvement in image and speech classification tasks for diverse scenarios such as millimeter wave and IEEE 802.11p vehicular channels.

Index Terms— Machine learning, channel estimation, artificial intelligence, wireless communications.

1. INTRODUCTION

Lately, the fifth generation (5G) wireless networks are very close to fully operational deployment. In order to design the physical layer in 5G networks, recently machine learning (ML) techniques have been introduced for the challenging design problems, e.g., resource management [1, 2], symbol detection [3], beamforming [4] and channel estimation [5]. While these methods have shown a great potential for system design to deal with data, hardware, and computational complexities, they have not been deployed in the current 5G architectures. Rather, the ML-based techniques are envisioned as the primary candidate to be considered in 6G network design [6–8].

In the upcoming 6G era, the wireless network will incorporate massive number of devices in the concept of internet of things (IoT), which provides ubiquitous sensing and computing capabilities to connect a broad range of things to the internet. These devices include mobile phones, connected vehicles, drones, smart sensors and industrial IoT devices. Hence, there will be a tremendous increase in the amount of data generated by these devices. The International telecommunications union (ITU) estimates that the global mobile traffic in 2030 will reach 5016 exabytes (EB), of which more than 70% will be image or video [9, 10]. Moreover, autonomous vehicles (AVs) are expected to generate approximately 20 TB/day/vehicle data, a part of which is transmitted to the road side infrastructure [11, 12]. To process and extract useful information from these images/videos, several ML techniques have emerged as a key enabling technology in the field of image classification, face/object/motion detection, target tracking, and speech recognition. These applications require huge amount of data to be processed and learned by an ML model for extracting the features from the raw data and provide a “meaning”. As a result, most of the processed data in 6G is expected to be generated or processed by ML algorithms.

The performance of the ML models depends on the data collected from the IoT devices in the network. However the wireless channel corrupts ML data (image/video/speech) during transmission and reduces the inference performance [3]. The corruption in the data may also occurs because of the sensor and network related errors. Thus, wireless channel acquisition is a crucial task in the ML applications. Furthermore, the 6G requirements, e.g., ultra-low latency and high mobility, make the channel acquisition process even more challenging [6]. To this end, several model-based [13, 14] and ML-based [4, 5, 15] channel state information (CSI) estimation techniques have been proposed for various communications standards in cellular (5G millimeter wave (mmWave)) and vehicular (IEEE 802.11p) networks. The model-based techniques have high energy and time consumption because they usually require tackling a high-dimensional optimization problem, especially for ultra-massive multiple-input multiple-output (UM-MIMO) configuration. On the other hand, the ML-based methods are data-driven, learn the features of the raw data, and provide robustness against noise and channel corruptions. Thus, it is instructive to leverage ML to address the uncertainty in channel estimates. However, implementation of ML techniques necessitates the task-oriented design of multiple ML models, each of which is dedicated to a different layer in the Open System Interconnection (OSI) communications model [7]. For instance, separate ML models for physical (e.g., channel estimation) and application (e.g., image recognition) layer tasks should be devised.

Motivated by the aforementioned two 6G facets — ML-related data and ML-based network design — we introduce implicit channel learning approach by combining the learning for both wireless channel and ML-related data. To this end, an ML model is trained...
with the dataset carrying the accompanied wireless channel such as path loss and correlation. Thus, the trained model implicitly learns the channel characteristics and makes accurate predictions even if the ML data are corrupted by the wireless channel. While there are a few studies on the implicit estimation of wireless channel for symbol detection [3] and active/passive beamforming [2, 16, 17], these methods require either a priori information such as user locations, or received pilot signals along with the optimization of the remaining system parameters, e.g., beamformers. Unlike these prior works, our proposed approach addresses both channel learning and application layer learning tasks through a single ML model, which is fed with the channel corrupted image/speech data. Thus, it is helpful for future 6G networks that are likely to heavily utilize ML-related data.

In the next section, we discuss common applications of ML in wireless communications. Next, we survey the state-of-the-art in channel estimation for various network architectures such as cellular and vehicular networks. In this context, we introduce our implicit channel learning approach and validate its performance on image and speech classification tasks under the effect of various wireless standards.

2. ML APPLICATIONS IN WIRELESS NETWORKS

In many contemporary ML applications, training and prediction of the ML model are carried out at a central parameter server (PS) of the network, whereas the data are generated at the network edge comprising the devices like mobile phones, connected vehicles, and IoT sensors. These ML models are composed of huge numbers of parameters. For instance, well-known ML models such as AlexNet, VGG (both for image classification), and GPT-3 (text recognition/translation) are comprised of 60 million (240 MB), 136 million (552 MB), and 170 billion (680 GB) learnable parameters, respectively [8, 9]. Storing these huge models in the edge devices is costly and inefficient. Instead, these models are stored at cloud PSs. The training stage is managed offline using a pre-collected training dataset. But the prediction-stage classification/recognition demands transmission of the generated image/video/speech to the PS because the ML models run short of storage at the IoT devices. Once the prediction is performed at the PS, its result is sent back to the IoT device. For example, Google’s image recognition technology Lens transmits captured images via a smart phone to the PS, where a pre-trained convolutional neural network (CNN) used for recognition/classification/translation brings up the information related to the objects in the captured image. Similarly, the AVs have on-board pre-trained learning models but still perform data transmission (reception) to (from) the cloud platforms for map generation, path planning, and forecasting [12].

Fig. 1 illustrates the impact of channel estimation on ML applications of image and speech classification for mmWave [13, 18] and vehicle-to-vehicle (V2V) [14] line-of-sight (LoS) links. We used MNIST and Speech Command datasets for image and speech classification tasks, respectively, each of which has 10 classes [9]. The image dataset includes the black-and-white handwritten digits whereas the speech dataset is composed of the spectrogram of audio signals. During training, clean (with equalized channel) datasets are used for both ML models while two different validation datasets are prepared with (clean) and without (corrupted) channel estimation to observe its effect on learning accuracy. We observe that the classification performance degrades significantly if the images/spectrograms are corrupted by the wireless channel without equalization for both tasks.

3. CHANNEL ESTIMATION TECHNIQUES

The model-based signal processing techniques need accurate mathematical modeling of the transmitted/received signals. However, to address the uncertainties and non-linearities imposed by channel equalization and hardware impairments, model-free ML techniques have become common in wireless communications [5, 8].

3.1. Model-based approaches

Channel estimation is an essential task for reliable communication [13]. However, it is more challenging in 6G architectures, wherein number of antennas in UM arrays is exceedingly huge [7, 13]. Furthermore, the multi-hop communications frameworks, such as vehicular networks [12, 15] and intelligent reflecting surface (IRS)-empowered systems [8, 16, 19] make channel estimation even more demanding. In particular, vehicular networks have highly dynamic channels arising from rapid vehicular mobility (up to 150 km/h in 6G). Then, variation in weather conditions (resulting in a path loss of ~4 dB over 0.1-0.3 THz) causes frequent drop-outs and hand-overs [12, 14]. As a result, there exists an inherent uncertainty stemming from the dynamics of the wireless channel in both network architectures.

3.2. ML-based approaches

Massive MIMO channel estimation via ML is investigated in [5], where a convolutional-only neural network (CoNN) is designed. The input of the CoNN is the tentative channel matrix that is computed via least squares (LS) method and the output is the channel matrix from multiple subcarriers. The CoNN exhibits improved channel estimation accuracy in terms of the mean-squared error (MSE). But its input data requires matrix inversion, which is computationally inefficient for large antenna systems. To reduce the complexity in preparing the input data, [4] devised a CNN approach for channel estimation by feeding the CNN with the received pilot signals. This
4. IMPLICIT CHANNEL LEARNING IN ML TASKS

Instead of designing dedicated learning models for both channel estimation and ML applications, both tasks could be performed jointly. Here, the ML model is trained on a training dataset while accounting for the imperfections or changes in the wireless channel. This allows the model to learn the corruptions in the ML data and perform the recognition tasks without performing channel estimation, leading to an overall reduction in the computational complexity and channel overhead. Also, without the channel estimation step, the ML model can only learn the statistical information about the channel, so the performance can be significantly worse than the method which uses the instantaneous CSI.

4.1. Data Collection and Training

The joint learning of the channel and the ML data requires preparing the training dataset for various channel conditions so that the ML model can extract the pattern in the input data and be robust against the corruptions of the wireless channel. Fig. 2 illustrates the processing chain of training data generation. Here, we generated three datasets each for both image and speech classification. The first set \( D_1 \) was clean while the second \( D_2 \) contained the corrupted data. Third set \( D_3 \) is a mix of 50% clean and 50% corrupted data. During data generation, the 5G and vehicular communication toolboxes of MATLAB are used, wherein the wireless channel statistics were randomly changed relying on Rician fading for each image transmission and the signal-to-noise-ratio (SNR) was set to 15 dB. After reception of the corrupted image data, it is quantized and resulting image is fed to the learning model, which has the knowledge of the labels and yields the classification result. Each corrupted dataset included 100 distinct corrupted copies of the clean dataset to provide robustness against imperfections. As a result, the number of samples in these three training datasets were 600,000, 6,000,000 and 3,300,000 (283, 700, 2,837,000 and 1,560,350) for MNIST (Speech Command) dataset, respectively. For the MNIST (Speech Command) dataset, there were two test datasets, each of which had a size of 10,000 (4,000) and they were generated separately from the training data.

4.2. Performance Evaluation

Table 1 and Table 2 show image and speech classification performance of the learning models, respectively, for various channel conditions such as mmWave, V2V, and multi-hop mmWave-V2V channels. The V2V channel was tested for multiple delay profiles corresponding to different scenarios, such as Rural LoS, Urban LoS/nLoS (non-LoS), and Highway LoS/nLoS [14, 15]. We observe that the learning performance was poor if there was a mismatch between the training and test datasets. In particular, the model trained on \( D_1 \) was unable to recognize the corrupted dataset, which includes the channel effects. On the other hand, the models trained on \( D_2 \) and \( D_3 \) were able to learn the corrupted ML data (image or speech) while implicitly learning the channel characteristics. Compared to the case conducted without channel estimation, they exhibit approximately 60% improvement for both image and speech classification. Furthermore, the model with \( D_2 \) provided slightly higher accuracy than the one possessing \( D_3 \) for the corrupted test data while the latter had a slight performance loss (approximately 1%) for clean test dataset. This is because the size of \( D_3 \) is smaller, including both clean and corrupted data. Nevertheless, \( D_3 \) presents satisfactory performance for both tasks with a smaller dataset. Consequently, this suggests that constructing the half of the dataset with channel effects can yield a reliable recognition performance as well as implicitly learning the channel effect on the ML data.

To compare the channel characteristics, the accuracy for mmWave-only channels degraded when combined with the V2V channel. This is explained by channel dynamics and the loss of beamforming gain, that mmWave usually leverages upon via multiple antennas.
The reliability of ML tasks aggravated in multi-hop scenario, i.e., mmWave-V2V for all delay profiles because of error propagation when the inaccurately received ML data in one vehicle was transmitted to the BS via the mmWave channel. Among all delay profiles, those with nLoS propagation had the most severe conditions leading to low classification accuracy. In particular, the ‘Highway nLoS’ showed the worst performance while ‘Rural LoS’ fared the best for all tasks in both V2V and mmWave-V2V channels.

Comparing the classification performance of both ML tasks, higher (~7%) accuracy was obtained for image classification than speech recognition. Note that both had the same number of classes. This is obvious because the image dataset had more distinguishable patterns (e.g., handwritten digits) whereas the features in the spectrogram of the audio signals were less prominent. When both training and test datasets were corrupted, the performance improvement in both tasks was within ballpark of each other due to the similarity. That is to say, they both get close to the clean dataset performance (i.e., 98.6% and 91.5% for, respectively, image and speech).

5. SUMMARY

We introduced implicit channel learning for ML applications in 6G networks. The proposed method jointly learns the features both in ML data and channel characteristics. By constructing a training dataset under the effect of various channel conditions in different propagation environments, the trained ML model becomes robust against the corruptions/imperfections. Compared to model-based methods, the ML-based techniques enhance the estimation performance and robustness against the channel dynamics. This is particularly helpful for highly dynamic channels at mmWave and THz, which also employ extremely large arrays.

Table 1. Image classification performance under different channel conditions.

| Learning Model: CNN with two convolutional layers (128@5 × 5 and 128@3 × 3) and a single fully connected layer (128 units) |
| Dataset: MNIST Handwritten Digits Dataset (28 × 28 × 60,000) |
| Classes: {0, 1, 2, 3, 4, 5, 6, 7, 8, 9} |
| Training Data: | Test Data: |
| D1 | Clean - Corrupted |
| Training Data: | Test Data: |
| D2 | Clean - Corrupted |
| Training Data: | Test Data: |
| D3 | Clean - Corrupted |

Dataset: 

| Class  | D1 | D2 | D3 |
|--------|----|----|----|
| Clean | 98.6% | 93.3% | 97.1% |
| mmWave | 95.1% | 97.0% |
| V2V - Rural LoS | 98.6% | 92.0% | 95.1% |
| V2V - Urban LoS | 98.6% | 97.8% | 87.9% |
| V2V - Urban nLoS | 98.6% | 91.7% | 87.7% |
| V2V - Highway LoS | 98.7% | 97.9% | 90.3% |
| V2V - Highway nLoS | 98.7% | 96.9% | 97.7% |
| mmWave-V2V - Rural LoS | 98.3% | 90.4% | 92.2% |
| mmWave-V2V - Urban LoS | 98.2% | 90.5% | 92.8% |
| mmWave-V2V - Urban nLoS | 97.7% | 83.8% | 88.1% |
| mmWave-V2V - Highway LoS | 98.4% | 90.4% | 90.3% |
| mmWave-V2V - Highway nLoS | 97.1% | 84.4% | 88.1% |

Table 2. Speech classification performance under different channel conditions.

| Learning Model: CNN with four convolutional layers (16@5 × 5, 16@3 × 3, 16@3 × 3 and 16@3 × 3) and a single fully connected layer (128 units) |
| Dataset: Speech Command Dataset (98 × 50 × 28, 370) |
| Classes: {yes, no, up, down, left, right, on, off, stop, go} |
| Training Data: |
| D1 | Test Data: |
| Clean - Corrupted |
| Training Data: |
| D2 | Test Data: |
| Clean - Corrupted |
| Training Data: |
| D3 | Test Data: |
| Clean - Corrupted |

Dataset: 

| Class  | D1 | D2 | D3 |
|--------|----|----|----|
| Clean | 91.5% | 29.6% | 15.2% |
| mmWave | 87.6% | 88.5% | 85.4% |
| V2V - Rural LoS | 91.9% | 77.1% | 80.0% |
| V2V - Urban LoS | 91.6% | 77.5% | 80.4% |
| V2V - Urban nLoS | 91.5% | 76.7% | 78.4% |
| V2V - Highway LoS | 91.5% | 79.0% | 80.2% |
| V2V - Highway nLoS | 91.5% | 75.3% | 81.7% |
| mmWave-V2V - Rural LoS | 91.5% | 85.1% | 82.6% |
| mmWave-V2V - Urban LoS | 91.6% | 84.2% | 82.3% |
| mmWave-V2V - Urban nLoS | 90.3% | 14.1% | 83.7% |
| mmWave-V2V - Highway LoS | 91.4% | 78.2% | 82.5% |
| mmWave-V2V - Highway nLoS | 91.4% | 26.3% | 85.1% |
6. REFERENCES

[1] L. Dai, R. Jiao, F. Adachi, H. V. Poor, and L. Hanzo, “Deep Learning for Wireless Communications: An Emerging Interdisciplinary Paradigm,” *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 133–139, 2020.

[2] A. Alkhateeb, S. Alex, P. Varkey, Y. Li, Q. Qu, and D. Tujkovic, “Deep learning coordinated beamforming for highly-mobile millimeter wave systems,” *IEEE Access*, vol. 6, pp. 37 328–37 348, 2018.

[3] H. Ye, G. Y. Li, B.-H. F. Juang, and K. Sivanesan, “Channel agnostic end-to-end learning based communication systems with conditional GAN,” in *IEEE Global Communications Conference Workshops*, 2018, pp. 1–5.

[4] A. M. Elbir, K. V. Mishra, M. R. B. Shankar, and B. Ottersten, “A Family of Deep Learning Architectures for Channel Estimation and Hybrid Beamforming in Multi-Carrier mm-Wave Massive MIMO,” IEEE Trans. Cognit. Commun. Networking, vol. 8, no. 2, pp. 642–656, Dec. 2021.

[5] P. Dong, H. Zhang, G. Y. Li, I. S. Gaspar, and N. Naderi-Alizadeh, “Deep CNN-based channel estimation for mmWave massive MIMO systems,” *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 989–1000, 2019.

[6] K. B. Letaief, Y. Shi, J. Lu, and J. Lu, “Edge artificial intelligence for 6G: Vision, enabling technologies, and applications,” *IEEE J. Sel. Areas Commun.*, vol. 40, no. 1, pp. 5–36, 2021.

[7] H. Yang, A. Alphones, Z. Xiong, D. Niyato, J. Zhao, and K. Wu, “Artificial-intelligence-enabled intelligent 6G networks,” *IEEE Network*, vol. 34, no. 6, pp. 272–280, 2020.

[8] A. M. Elbir, A. K. Papazafeiropoulos, and S. Chatzinotas, “Federated learning for physical layer design,” *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 81–87, 2021.

[9] R. Mayer and H.-A. Jacobsen, “Scalable deep learning on distributed infrastructures: Challenges, techniques, and tools,” *ACM Comput. Surv.*, vol. 53, no. 1, pp. 1–37, 2020.

[10] A. M. Elbir and K. V. Mishra, “Cognitive Learning-Aided Multi-Antenna Communications,” *IEEE Wireless Commun.*, vol. 29, no. 6, pp. 136–143, May 2022.

[11] A. M. Elbir, B. Soner, S. Çöleri, D. Gündüz, and M. Bennis, “Federated Learning in Vehicular Networks,” in *2022 IEEE International Mediterranean Conference on Communications and Networking (MeditCom)*. IEEE, Sep. 2022, pp. 72–77.

[12] J. Posner, L. Tseng, M. Aloqaily, and Y. Jararweh, “Federated learning in vehicular networks: Opportunities and solutions,” *IEEE Network*, vol. 35, no. 2, pp. 152–159, 2021.

[13] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, “Channel estimation and hybrid precoding for millimeter wave cellular systems,” *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, 2014.