

Channel State Information Prediction in 5G Wireless Communications: A Deep Learning Approach

Changqing Luo, Jinlong Ji, Qianlong Wang, Lixing Yu, and Pan Li

Abstract—Channel state information (CSI) estimation is one of the most fundamental problems in wireless communications, the success of which can boost the performance of wireless communications significantly. Various kinds of technologies have been proposed to address this problem. However, they are usually expensive in throughput used for transmitting pilot signals. This can be even worse in 5G wireless communications. Besides the high throughput cost, employing massive MIMO and millimeter-Wave (mmWave) communications incurs expensive computational cost in conducting CSI estimation. In contrast, we notice that CSI essentially reflects electromagnetic wave propagations which can be affected by many features like location, time, weather, humidity, etc. Therefore, CSI estimation for a specific location should be conducted based on the historical CSI and other features' data. In this paper, we employ a deep learning technique, i.e., recurrent neural networks (RNN), to predict the CSI values. We conduct extensive experiments to evaluate the performance of the proposed model using the data collected from our own constructed testbeds. The results show that the RNN technique is able to perform accurate CSI estimation in wireless communications.

Index Terms—Channel state estimation, 5G wireless communications, deep learning

1 INTRODUCTION

Channel state information (CSI) is one of the most fundamental concepts in wireless communications. It refers to known channel properties of a radio communication link. Specifically, CSI can characterize the combined effect of path loss, scattering, diffraction, fading, shadowing, etc., when a signal propagates from a transmitter to its corresponding receiver. As a result, this information can tell us whether a radio communication link is in a good or bad state.

Obtaining accurate CSI is of paramount importance in guaranteeing the performance of radio communication links in wireless communication systems. The deployment of physical layer parameters mainly depends on the this information in wireless communications. For example, when a channel is in a bad condition, low-order modulation schemes need to be employed to combat the poor channel and reduce bit error rate, while once a channel is in a good state, high-order modulation schemes need to be deployed for obtaining high data transmission rate. If CSI is not accurate, a wrong modulation scheme will be employed, which incurs high data error rate or low data transmission rate. Moreover, it is also known that the control of radio resources [1] and interference management [2] in wireless communications are highly related to CSI. Thus, to guarantee the performance of radio communication links, we need to find accurate knowledge about CSI of a channel for wireless communications.

To find accurate CSI, researchers have proposed many methods for CSI acquisition that is composed of pilot design [3] and channel estimation. For the former, to obtain CSI at a radio receiver, the receiver needs to first send some signals to its transmitter and receive the feedback. Some previous studies [4]

consider to send pilot signals to its transmitter and receive the feedback. This, however, results in high CSI acquisition delay. To address this problem, some works exploit piggybacking technologies to obtain feedback from packets sent from the transmitter [5]. These methods introduce large amount of overhead in wireless communication systems since they require each packet to carry the side information. For the latter, channel estimation is conducted to obtain CSI after obtaining the feedback. So far, many methods have been developed to conduct channel estimation, such as maximum likelihood (ML) estimation [6], least-square (LS) estimation [7], minimum mean square error (MMSE) estimation [8]. For example, Du *et al.* [9] consider to perform ML based channel estimation for macro-cellular OFDM uplinks in time-varying channels. The channel state is determined by estimating the unknown channel parameters. Karami [10] considers to employ an LS algorithm to perform MIMO channel estimation, and shows the tracking performance. Ma *et al.* [11] propose a linear MMSE method to estimate CSI of individual channels, and especially the initialization point for the iterative linear MMSE is obtained by using an LS channel estimator. We notice that by employing these algorithms, we need to perform matrix operations like matrix multiplication, matrix inverse, eigenvalue decomposition, and singular value decomposition. Performing such matrix operations requires high computational complexity. To sum up, we notice that such works for signal design and channel estimation usually have large amount of overhead and high computational complexity, which makes them only suitable for small-scale wireless communication systems.

On the other hand, current cellular communication systems are now evolving into the next-generation (5G) wireless communication systems. Past few years have witnessed the exploding growth in mobile devices. To serve the massive amount of mobile devices, 5G wireless communication systems will be large-scale. Moreover, we observe the exponential increase in mobile data

C. Luo, J. Ji, Q. Wang, L. Yu, and P. Li are with Department of Electrical Engineering and Computer Science, Case Western Reserve University, Cleveland, OH 44106. Email: {cxl881, jxj405, qxw204, lxy257, lipan}@case.edu.

traffic in wireless communication systems due to various emerging mobile applications like mobile health [?], smart cities [?], and internet of things (IoT) [?] are emerging. According to Ericsson mobility report 2017 [?], worldwide total mobile data traffic was already up to 8.8 ExaBytes in 2016 and this figure grew by 70% in 2017. To accomodate so massive amount of traffic data in future 5G wireless communication systems, massive multiple-input multiple-output (MIMO) and millimeter-Wave (mmWave) communication have been considered as two of ideal candidate techniques. Massive MIMO exploits spatial domain resources for offering diversity gain, multiplexing gain, and power gain [?] and mmWave communication enables 5G wireless communication systems to utilize the spectrum ranging from 3 GHz to 300 GHz [?].

Due to the large size and the newly-developed technologies, conventional CSI acquisition methods are faced with a number of challenges when they are used in 5G wireless communication systems. First, the overhead incurred by transmitting pilot signals is very high. The reasons are fold. One is due to serving the great number of mobile devices. Each mobile device needs to obtain pilot signals before performing CSI estimation. The other is due to employing massive MIMO antennas. The amount of CSI feedback needs to be scaled with the number of antennas to control the quantization error. This hence incurs more overhead to be transmitted over 5G wireless communication systems. Second, the computing time used for conducting CSI estimation is very long. 5G wireless communication systems have much more channels than current cellular communication systems, which makes the channel estimation problem much more complicated. When Conventional CSI estimation methods are used to perform CSI estimation, they require performing large-scale matrix operations with high computational complexity, which incurs the formidable challenging for the mobile device with very limited computational resources. Such mobile devices need to spend a very long time to obtain the estimated CSI. This is impractical for 5G wireless communication systems because CSI must be refreshed at a rate depending on the correlation time of the channel. Therefore, we need to find a very efficient way to perform CSI estimation in 5G wireless communications.

In this paper, we propose an online

We summarize our major contributions as follows:

- 1) We introduce a three-layer architecture for large-scale 5G wireless communications.
- 2) We propose to dynamically control transmission power for improving NLOS transmission performance in large-scale 5G wireless communications.
- 3) We explore UE association with MBSs/SBSs and radio resources allocation, and formulate the maximization problem of UEs' sum-rate under the constraints of radio resources and UEs' QoS requirements.
- 4) To efficiently solving the formulated problem, we propose a deep Q-network method. In particular, we apply a convolutional neural network (CNN) to perform the Q-function estimation offline, and based on the estimation results, conduct the deep Q-learning online to find the control strategy (i.e. UE association with MBS/SBSs and radio resources allocation). Due to estimating the Q-function offline, the online control is very efficient.
- 5) We conduct extensive simulations to validate the efficacy of the proposed scheme and simulation results show

that our proposed scheme can improve the performance significantly.

The rest of this paper is organized as follows. We first briefly introduce the preliminary knowledge about the radio signal propagation in Section 2. Then, we describe the proposed CSI prediction scheme in detail, including the learning framework, the dataset construction, and the proposed learning method in Section 3. Afterwards, we present the experiments results thoroughly in Section 4. Subsequently, we present the most related works in Section 5, and finally conclude this paper in Section 6.

2 PRELIMINARIES

In this section, we briefly introduce radio signal propagation. In 5G wireless communication systems, radio signals are transmitted over transmission media, i.e., air. A typical radio signal propagation system is shown in Fig. ?? In this figure, we can observe that a transmitter transmits a signal x over air to a receiver who receives a signal y . Specifically, the signal is transmitted through an antenna which converts the radio signal into an electromagnetic wave that is propagated over free space. The signal transmission can be expressed as follows:

$$y = \sqrt{P_t h} x + n, \quad (1)$$

where P_t is the transmission power at the transmitter, h is the channel gain (i.e., CSI), and n is the noise received by the receiver.

In practice, electromagnetic wave propagates over the air without any protection, a radio signal is prone to suffering from the attenuation of transmission power. The CSI is used to represent the attenuation of a radio signal over the air, i.e., the combined effect of the path loss, scattering, fading, etc. It is well-known that radio channels suffer from path loss unavoidably due to the dissipation of the power radiated by the transmitter. In free space, there is no obstacles between a transmitter and a receiver and thus a transmitter can perform line-of-sight (LOS) signal transmissions to transmit radio signals to a receiver that is d meters away. Based on [?], we can have an approximation expression on the path loss as follows:

$$Pl = 10 \log_{10} \frac{P_t}{P_r} = -10 \log_{10} \frac{G_t \lambda^2}{(4\pi d)^2}, \quad (2)$$

where Pl is the path loss in decibel (dB), P_r is the receiving power, G_t is the product of the transmit and receive antenna field radiation patterns of LOS transmissions, and λ is the operating wavelength that can be calculated by $\lambda = 3 \times 10^8 / f$ (here, f is the frequency). Hence, we can see that the path loss depends on the distance between a transmitter and its receiver for transmitting a radio signal over a specific channel through a specific antenna.

Moreover, scattering and fading also have impact on the CSI. Scattering is natural in wireless communications due to the radiation of electromagnetic wave transmission in the medium and the obstacles. When the density of the air and humidity are different, the impact of the scattering on the CSI is also different. On the other hand, the fading of a wireless communication link may either be because of multi-path propagation, weather like rain, or shadowing from obstacles. Due to the changing rate of the magnitude and phase electromagnetic waves, we can usually have slow fading and fast fading. Besides, the fading is also related to the frequency band. Signals transmitted over some frequency bands may suffer from severe performance degradation, which refers to frequency selective fading. Moreover, some frequency

bands have some other features. For example, as reported by [?], the atmosphere absorbs the power of electromagnetic waves and the absorption peaks occur at 24 GHz and 60 GHz when radio signals travel through the atmosphere and there is stable loss window between two peaks.

To estimate the CSI, we need to send pilot signals and get feedback. Then, based on the knowledge of the transmitted and received signals, we can conduct CSI estimation. Let p_1, \dots, p_N denote pilot sequence, and we can have

$$y_i = H p_i + n_i, \quad (3)$$

where H is the channel matrix. By combining N received pilot signals, we can have

$$Y = HP + N. \quad (4)$$

H needs to be recovered from the knowledge Y and P . To recover H , researchers have developed several methods like ML, MMSE, and LS.

3 ONLINE CSI PREDICTION SCHEME FOR 5G WIRELESS COMMUNICATIONS

In this section, we describe the online CSI prediction framework, the dataset construction, and the proposed deep learning network, respectively.

3.1 Online CSI Prediction Framework

From Section 2, we notice that the CSI of a radio communication link is related to path loss, specific frequency band characteristics, weather, atmosphere, and obstacles, etc. One question raises: is it possible for us to employ a certain machine learning technique for analyzing the dataset constructed by all related features and the measured CSI and find the relationship between CSI and all related features? If yes, this can definitely help predict CSI efficiently for a specific frequency band at a place.

Based on the analysis of all the features related to the CSI, we can note that such features have some certain patterns. Specifically, a specific frequency band naturally has particular electromagnetic propagation characteristics that does not change along time and locations, which implies that different frequency bands have different CSI at the same place and at the same time. As a result, frequency bands have some certain patterns in the frequency domain. Moreover, the atmosphere factors like water vapor, molecules of oxygen, and other gaseous atmospheric components are in fact related to the time domain. For example, the measured values of such atmosphere factors in the daytime are different from that during the night. Thus, such atmosphere factors tend to have certain patterns in the time domain. Furthermore, some obstacles like buildings and trees are usually static and they are always at some specific locations. Some other obstacles are dynamic, such as walking people and cars. For example, a crowd of students walking during the daytime of weekdays in a campus and only a few students there during the evening and weekends. Thus, such dynamic obstacles tend to have some certain patterns as well. As a result, obstacles have some certain patterns in the space domain. In addition, different degrees of rain can have different effect on electromagnetic wave propagations, which is helpful for us to predict CSI in 5G wireless communications. To sum up, these features related to the CSI of a radio communication link have some certain patterns that we can extract. We may take advantage of such extracted patterns to perform CSI estimation

for a specific channel at a place during a time in 5G wireless communication systems.

To efficiently find CSI for a mobile device at a place during a time period, we propose an Online CSI prEdiction scheme for 5G wireless communicAtions, i.e., called OCEAN. The detailed framework of OCEAN is shown in Fig. ???. Specifically, OCEAN first employs a learning framework for analyzing historic data and extracting patterns from them, i.e., training the model. Then, when a CSI prediction request arrives, OCEAN collects real-time data of all features, and inputs these data into the learning framework to perform online CSI prediction and output the CSI for different frequency bands. In particular, we propose a new learning framework for accurate CSI prediction. This learning framework is the combination of a convolutional neural network (CNN) and long short term memory (LSTM) network (i.e., a variation of recurrent neural networks). This is because of the certain patterns existing in frequency, time, and space domains. In the following two sub-sections, we describe two important parts of our proposed OCEAN, i.e., dataset construction and learning framework.

3.2 Dataset Construction

In our proposed OCEAN, we consider all related features, including frequency band, location, time, temperature, humidity, and weather. In the following, we describe these features used in OCEAN one by one.

Frequency band: 5G wireless communications utilize the spectrum ranging from 3 GHz to 300 GHz. In practical 5G wireless communications, the spectrum is divided into a number of channels, each of which corresponds to a frequency band. Thus, in OCEAN, there are a set of channels that are denoted by $\mathcal{F} = \{f_1, f_2, \dots, f_N\}$. In practice, at a time, a mobile device experiences different CSI values for different frequency bands.

Location: In 5G wireless communication, a base station covers an area. At different locations, mobile devices have different CSI even for the same frequency band. In OCEAN, the coverage area is initially divided into several sub-regions based on the distance to the base station. However, due to the obstacles' influence, mobile users in the same sub-region may have very different CSI values. To solve this problem, we first calculate the mean *mean* and root-mean-square deviation (RMSD) *RMSD* for a sub-region. If *RMSD* is larger than a predefined value ϵ , we consider to separate a sub-region into three: one sub-region contains all mobile devices with the values within $[mean - RMSD, mean + RMSD]$, the second one contains all mobile devices with the values less than $mean - RMSD$, and the third one contains all mobile devices with the values larger than $mean + RMSD$. Following this way, we check all the initial sub-regions. Then, we merge the sub-regions whose RMSD is less than ϵ . Thus, we can finally obtain the number of sub-regions that are denoted by $\mathcal{L} = \{l_1, l_2, \dots, l_M\}$.

Time: In practice, the dense of the atmosphere is different in different season and also in different period during a time. As a result, in OCEAN, we consider the format of the time as "xx:xx:xx, month/day/year". To reduce the amount of similar data, we consider to collect data every 3 min.

Temperature: The different temperature can lead to the different dense of the atmosphere, hence affecting the effect of scattering and fading of radio signal propagations. In OCEAN, we consider to collect the real-time temperature data and the temperature is denoted as $\{t_1, t_2, \dots, t_P\}$.

Humidity: The humidity also has impact on the effect of scattering and fading. In OCEAN, we consider to collect the real-time humidity data and the humidity is denoted as $\{u_1, u_2, \dots, u_H\}$.

Weather: As stated before, the weather has very important impact on the CSI. Particularly, rain can absorb the power of radio signals transmitted over some specific frequency bands. In OCEAN, the set of the weather is $\{\text{sun, sun with cloud, cloud, light rain, medium rain, heavy rain, light snow, medium snow, heavy snow}\}$.

3.3 The Proposed Learning Framework

Our proposed scheme, online CSI prediction scheme for 5G wireless communications (OCEAN), is an online neural-based end-to-end system to predict CSI value in various of scenarios. The architecture of OCEAN system is structured as Figure X. The whole system is divided into four stages. We construct CSI information maps from raw CSI data and other side information. The standard convolutional layers in first stage extract frequency representative (FR) vectors from CSI information maps of different frequency bands. The following 1D convolutional neural network compresses FRV of multiple bands into one state representative (SR) vector. In stage 3, the SR vectors pass through the LSTM layers to predict future SR vectors. At the output of the LSTM layers, we reconstruct CSI information maps from the SR vectors. To automatically adjust the system according to practical scenario, we equip our OCEAN system with online learning mechanism, which makes the system more efficient and stable in practical scenario. In this section, first we introduce CSI information maps, the training dataset used for OCEAN system. Then we describe the details of each stage of our model, like frequency representative vector extraction, state representative vector extraction, state vector prediction and reconstruction of predicted CSI information maps. Finally, we present the online learning mechanism used in OCEAN system.

3.3.1 CSI Information Image

To use the proposed learning framework, the input data need to be preprocessed before feeding them into the learning framework. This is because the first part, i.e., the CNN network, is originally developed for processing image data. As a result, to make the proposed learning framework work, we need to organize the input data.

To train color image data by a CNN network, the input data is the pixel that is the smallest addressable element in an image. Each image is composed of pixels that is represented by a series of code. A channel is the grayscale image with the same size as the color image.

Inspired by the representation of color images, we organize the raw data in the similar way. Specifically, the raw data is segmented into small cells, each of which corresponds to a pixel in an image. For each frequency band, the data of the CSI and other features are formed the pixels.

Inspired by the representation of digital color image, we construct CSI information maps from raw CSI data and other side information. Color digital images are made of pixels, and pixels are made of combinations of primary colors represented by a series of code. A **pixel** is the smallest addressable element in the image. A **channel** is the grayscale image of the same size as a color image, made of just one of these primary colors. Similarly, we segment the map of a certain place into several small cells, which

is "CSI pixel". For each frequency band, we collect CSI value and side information in these CSI pixels. The maps, created of CSI value data of same frequency, or same types of side information, are like "CSI information channels". The CSI information maps are the combination of CSI information channels.

The training dataset is the collection of CSI information maps. At each time slot, we collect required data and build the CSI maps X_i . As time goes by, we obtain a CSI information maps sequence (X_1, X_2, \dots, X_i) .

3.3.2 Frequency Representative Vector Extraction

The CSI information maps contains details of raw information of CSI values and other side information data. For classic machine learning system, we should manually extract features from these data, and select suitable features, which will deeply influence future CSI values. However, the feature extraction and selection processes are heavily depends on expertise on CSI. In addition, the manually extraction and selection are effort-consuming tasks. To address these issues, we utilize the advance of neural-based learning system to extract representative vectors from the CSI information maps.

As show in stage 1 of Figure X, the convolutional neural networks (CNNs) extract frequency representative vectors from CSI information maps. (Here, the vectors obtained by CNN are not only frequency representative vectors. In fact, several vectors represent the side environment information, such as *, *. We ignore the differences between these two types of vectors, and call all of them Frequency Representative Vectors) First, we decompose a CSI information map into multiple CSI information channel

$$\alpha = \text{softmax} \left(\mathbf{w}^T \tanh(\mathbf{W}_h \mathbf{H}_a) \right) \quad (5)$$

$$(6)$$

For normal CNNs, we use a number of convolutional filters to process these channels, but for clarity we will explain the CNN networks with one filter. Let w_R be a convolutional filter which we apply to a window of pixels in the channel to generate a feature. A feature C for a window of pixels (P_1, P_2) is given as follows:

$$\alpha = \text{softmax} \left(\mathbf{w}^T \tanh(\mathbf{W}_h \mathbf{H}_a) \right) \quad (7)$$

$$(8)$$

where b_v is a bias term and X is the activation function. A feature map C is a collection of features computed from all windows of pixels:

$$\alpha = \text{softmax} \left(\mathbf{w}^T \tanh(\mathbf{W}_h \mathbf{H}_a) \right) \quad (9)$$

$$(10)$$

To capture the most salient features in c , we apply a max-over-time pooling operation (Collobert et al., 2011), yielding a scalar:

$$\alpha = \text{softmax} \left(\mathbf{w}^T \tanh(\mathbf{W}_h \mathbf{H}_a) \right) \quad (11)$$

$$(12)$$

In the end of the stage 1, we combine all outputs from convolutional filters together, and flatten them into a vector, which is the extracted Frequency Representative Vector.

3.3.3 State Representative Vector Extraction

After Stage 1, we obtain FR vectors of all frequency bands and side information. Concatenating these FR vectors together, we build a state representative matrix of a certain time slot. We can use the matrix sequence as input to pass it through recurrent neural networks to capture the temporal features of the sequence. However, sequentially concatenating vectors cannot imply the inference among frequency bands and side information. Meanwhile, the size of dimension of inputs determines the number of parameters required to optimize. If the input size is too large, there are a huge size of parameters need to train, which will spend an unacceptable time, and the system may not converge for lacking enough training samples.

To solve these problems, we use CNNs to compress the state representative matrix into a state representative (SR) vector. The process is similar to FR vectors extraction, but for SR vector extraction, the convolutional computation is only along the one dimension as shown in Figure X1. The output of this CNN is a state representative vector:

$$\alpha = \text{softmax}(\mathbf{w}^T \tanh(\mathbf{W}_h \mathbf{H}_a)) \quad (13)$$

$$(14)$$

3.3.4 State Vector Prediction

For OCEAN system, we aim at predicting the future CSI state of the whole space based on previous and current CSI state. From previous stage, we accumulate the captured state representative vectors to form a state vector sequence. Due to the superior performance of recurrent neural networks in different sequence learning tasks, we adopt recurrent neural network to predict future state vector, specifically we choose LSTM-based RNN.

Base on the Equation X, the output of LSTM network layers can be written as:

$$\alpha = \text{softmax}(\mathbf{w}^T \tanh(\mathbf{W}_h \mathbf{H}_a)) \quad (15)$$

$$(16)$$

Then, a softmax layer is followed to transform XX to predicted state vector.

$$\alpha = \text{softmax}(\mathbf{w}^T \tanh(\mathbf{W}_h \mathbf{H}_a)) \quad (17)$$

$$(18)$$

3.3.5 Reconstruction of Predicted CSI Information Maps

The output of previous stage is the predicted state vector. However, the goal of our design is to predict the real multi-spectrum CSI status of a specific space, which is represented by CSI information maps. Only obtaining status vector is far from our objective.

**We segment the state vector into n part. The length of segments should equal to the size of the channel in the CSI information map. **

3.3.6 Online Learning Mechanism

Although we consider plenty of side information, like temperature, humidity, obstacles etc. in our system design, there still are some factors which may lead to abrupt change of CSI status. To address this issue, we equip our system with online learning mechanism, which can capture any inference to CSI status and make the system more stable in real scenario.

Set Threshold

4 EXPERIMENTS

In this section, we evaluate the performance of the proposed CSI prediction scheme through four typical case studies, i.e., **Case I:** free space environment, **Case II:** outdoor environment, **Case III:** a work room, and **Case IV:** a building. We believe that considering the CSI prediction in these scenarios is very helpful for validating our proposed CSI prediction scheme. In what follows, we describe the data collection and preprocessing, RNN-CNN network settings, and experiment results, respectively.

4.1 Data Preparation

We collect the CSI data in the considered four scenarios. To collect the data, we use a Dell laptop equipped with an Intel Wireless Link 5300 network interface card as the mobile device and configure a TP-Link WiFi router as an access point. The mobile device sends successive packets to the WiFi router, and extracts the CSI from the received packets.

For **Case I:**, we consider a free space environment and collect data in a scenario as shown in Fig. ???. In this scenario, we place the WiFi router at the window on the second floor and the wireless router is xx meters away from the ground. Moreover, we move the laptop around the parking lot and collect CSI data at each point. In addition, there is no obstacles in the parking lot and the information transmission between the wireless router and the laptop is an LOS transmission. Thus, this is approximately a free space environment.

For **Case II:**, we consider an outdoor environment as shown in Fig. ???. We move the laptop around out of a building where there are trees and people walking around on the ground. Moreover, the laptop communicates with a wireless router which is placed at the window on the fourth floor and is high off the ground for xx meters. This is a typical outdoor environment in which wireless communications often occur.

For **Case III:**, we consider a typical indoor environment, i.e., a work room in a building as shown in Fig. ???. This is a laboratory in a building and there are several cubics and desks in this room. As shown in the figure, we place the wireless router on the desk. To collect the CSI data, we move the laptop in this room and stay each collecting point for a while. This indoor environment is also a typical scenario because wireless communications often occur at workplaces as well.

For **Case IV:**, we consider another typical indoor environment, i.e., a building as shown in Fig. ???. In this building, there are 5 floors and we collect the CSI data at one of the floor. Specifically, we place the wireless router on the desk of a room and move the laptop on this floor. To collect the CSI data, we place the laptop in several rooms, and especially the laptop stays at each position for a period. This scenario is similar to that of **Case III**. The major difference is that we consider the impact of the walls on the CSI in this case.

Moreover, we collect the weather data from xxx.

4.2 RNN-CNN Network Settings

In our experiments, we build our RNN-CNN network with multi-dimension input data. The output of the RNN network has 4 neurons. The activation function is set by default to the sigmoid function and make a single value prediction. The "look back window" of the system is set to 1 and 5, which is the number of most recent data that we consider as input to predict the next time slot variables.

Moreover, we also set up a CNN used for estimating the Q-function. There are three convolutional layers, two pooling layer, and two fully connected layers. The values of the initial weight vector are set as random values.

4.3 Experiment Results and Discussions

To measure the performance of the proposed CSI prediction scheme, we adopt the mean squared derivation (MSD) to assess the derivation between the prediction values and the measured ones. To compare the better performance of our proposed CSI prediction scheme, we conduct the experiments by only using the CNN network to predict the CSI. In the following, we show the performance of our proposed CSI prediction scheme in the considered four case studies.

4.3.1 Case Study I: A Free Space Environment

Fig. ?? shows the prediction results along the time. Specifically, we compare the training and testing results with the real CSI values collected in this case study. From this figure, we can find that the training and testing results are very close to the real CSI values. This implies that our proposed scheme can work well in predicting CSI data.

Fig. ?? illustrates the prediction precision of our proposed scheme and compares the performance with other two schemes. Specifically, the MSD values of our proposed scheme, LSTM, and MLP are xx, xx, and xx, respectively. The proposed scheme has the lowest MSD value. Moreover, the LSTM scheme is better than the MLP scheme. Therefore, both observations give us a sight that we need to design a sequence-related scheme to conduct CSI prediction.

Fig. ?? also demonstrates the performance of the proposed scheme. Specifically, we compare the prediction results with that calculated by using Eq. (??). From this figure, we can see that the curve of the obtained prediction results is close to that of the calculated results. This also validates the efficacy of our proposed scheme. Moreover, Specifically, the MSD values of our proposed scheme, LSTM, and MLP are xx, xx, and xx, respectively. The proposed scheme has the lowest MSD value. Moreover, the LSTM scheme is better than the MLP scheme. Therefore, both observations give us a sight that we need to design a sequence-related scheme to conduct CSI prediction.

4.3.2 Case Study II: Outdoor Environment with Obstructions

Fig. ?? shows the prediction results along the time. Specifically, we compare the training and testing results with the real CSI values collected in this case study. From this figure, we can find that the training and testing results are very close to the real CSI values. This implies that our proposed scheme can work well in predicting CSI data.

Fig. ?? illustrates the prediction precision of our proposed scheme and compares the performance with other two schemes. Specifically, the MSD values of our proposed scheme, LSTM, and MLP are xx, xx, and xx, respectively. The proposed scheme has the lowest MSD value. Moreover, the LSTM scheme is better than the MLP scheme. Therefore, both observations give us a sight that we need to design a sequence-related scheme to conduct CSI prediction.

4.3.3 Case Study III: A room of An Indoor Environment

Fig. ?? shows the prediction results along the time. Specifically, we compare the training and testing results with the real CSI values collected in this case study. From this figure, we can find that the training and testing results are very close to the real CSI values. This implies that our proposed scheme can work well in predicting CSI data.

Fig. ?? illustrates the prediction precision of our proposed scheme and compares the performance with other two schemes. Specifically, the MSD values of our proposed scheme, LSTM, and MLP are xx, xx, and xx, respectively. The proposed scheme has the lowest MSD value. Moreover, the LSTM scheme is better than the MLP scheme. Therefore, both observations give us a sight that we need to design a sequence-related scheme to conduct CSI prediction.

4.3.4 Case Study IV: A building of An Indoor Environment

Fig. ?? shows the prediction results along the time. Specifically, we compare the training and testing results with the real CSI values collected in this case study. From this figure, we can find that the training and testing results are very close to the real CSI values. This implies that our proposed scheme can work well in predicting CSI data.

Fig. ?? illustrates the prediction precision of our proposed scheme and compares the performance with other two schemes. Specifically, the MSD values of our proposed scheme, LSTM, and MLP are xx, xx, and xx, respectively. The proposed scheme has the lowest MSD value. Moreover, the LSTM scheme is better than the MLP scheme. Therefore, both observations give us a sight that we need to design a sequence-related scheme to conduct CSI prediction.

To sum up, the above experiment results have clearly shown the effectiveness of our proposed online CSI prediction scheme. In particular, we can see that our proposed scheme is able to achieve very low MSD values. Moreover, comparing with other two schemes, our proposed scheme has better performance and especially the one without considering the temporal dimension has the poorest performance, which gives us a sight that the CSI prediction doesnot only depend on the spatial but also the temporal.

5 RELATED WORKS

So far, previous studies have examined how to apply machine learning techniques into 5G wireless communications and networks. Specifically, these works have studies some important issues in 5G wireless communications and networks by exploring machine learning methods, such as some mobile traffic prediction, wireless network optimizations, and user behaviours. We summarize them in the following.

Researchers have investigated the mobile traffic forecasting problem in wireless networks. Wang *et al.* [?] modeled spatial-temporal mobile traffic data based on the existing dataset and proposed to employ deep learning approaches for the traffic prediction. Sciancalepore *et al.* [?] considered network slicing traffic forecasting and developed a Holt-Winters theory based forecasting solution to predict future traffic levels per network slice. Han *et al.* [?] also analyzed the mobile traffic load for a customer and predicted busy hour for each month by exploiting the support vector machine (SVM) technique.

Another research line is to optimize wireless network design by employing machine learning techniques. Kato *et al.* [?] characterized mobile traffic in wireless heterogeneous networks and proposed a deep learning based method to control mobile traffic flowing among multiple wireless access networks. Xu *et al.* [?] considered the problem of radio resource allocation in cloud radio access networks for 5G wireless communications and developed a radio resource allocation strategy based on a deep reinforcement learning that is used to approximate the action-value function. He *et al.* [?], [?] proposed to utilize a deep reinforcement learning technique for jointly optimizing both caching and interference alignment in 5G wireless communications. Liu *et al.* [?] considered to identify the link usage in the flow constrained optimization problem in 5G wireless networks by using the deep learning method, in order to reduce the scale of the considered optimization problem.

Some other previous works considered to study some other important issues in 5G wireless networks by using machine learning techniques, such as user behaviors and quality of experience (QoE). Parwez *et al.* [?] utilized wireless network data, i.e., call detail record, to identify anomalous behavior of a wireless network. Lin *et al.* [?] applied a machine learning method to predict users' QoE by considering users' features like user numbers and CSI experienced by a user.

From the above analysis, we can notice that machine learning techniques are now considered as one of promising tools in designing 5G wireless networks and improving the wireless network performance. However, few of them apply such techniques into predicting the CSI of radio signal propagation in wireless networks.

6 CONCLUSIONS

In this paper, we have studied the problem of enhancing NLOS transmission performance in large-scale 5G wireless communications. To improve NLOS transmission performance, we propose to dynamically control transmission power. In particular, we explore the control of UE association with MBSs/SBSs and radio resources allocation, and formulate the maximization problem of UEs' sum-rate under the constraints of radio resources and UEs' QoS requirements.

REFERENCES