A Blind CSI Prediction Method Based on Deep Learning for V2I Millimeter-Wave Channel

Jingxiang Yang, Liyan Li, and Min-Jian Zhao
College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China
Email: {21831136, liyan_li, mjzhao}@zju.edu.cn

***Abstract*—With the development of the Internet of vehicles and 5G, there emerge more and more challenging application scenarios with fast time-varying channels and high mobility nodes, such as high speed trains environment and vehicle-to-infrastructure (V2I) communication in highway. To support the reliable vehicular communication and mobile edge computing (MEC), it is important to obtain the future channel state information (CSI), which can help optimize system transmission scheme. In this paper, we propose an efficient blind CSI prediction model, called BCPMN. We first reshape the sampled signal into a specific 2-dimensional matrix. Then we propose a learning framework contains of convolutional neural network (CNN), long short-term memory (LSTM) network and fully connected layers. To validate the proposed model, we conduct extensive experiment in three modulation modes. The results show that the BCPMN achieves highly accurate signal-to-noise ratio (SNR) prediction in the fast changing channel model with different modulation modes. In particular, the proposed model can obtain better performance than other methods, and can achieve better performance than other methods without the payload cost of pilot.**

***Index Terms*—Vehicular communication, V2I, CSI, BCPMN, SNR prediction.**

I. Introduction

With the development of 5th generation wireless systems (5G), more and more vehicular communication scenarios are beginning to apply 5G [1] such as vehicle-to-vehicle (V2V) communication [2], vehicle-to-infrastructure (V2I) communication [3] and high speed train (HST) communication [4]–[6]. Due to the high sensitivity of millimeter-wave (mm-wave) wireless system to channel quality and environmental conditions [7], providing more effective and reliable communication services for 5G vehicular communication system is becoming more important. Recently, mobile edge computing (MEC) [8] and adaptive transmission scheme [9]–[11] is widely studied. Under low speed environments, Libo Jiao [8] provided an MEC system using imperfect channel state information (CSI) to achieve efficient resource allocation and Gavin Holland [11] proposed an adaptive rate and modulation scheme based on signal-to-noise ratio (SNR) to improve high network throughput, which illustrates that CSI, such as SNR, is highly essential for wireless communication systems to optimize the performance.

D. R. Pauluzzi [12] compared some classic SNR estimation algorithms and summarized several excellent and easy to implement estimators such as maximum-likelihood (ML) estimator, squared signal-to-noise variance (SNV) estimator and second- and fourth-order moments (M_2M_4) estimator. The SNV estimator is actually a special case of the ML SNR estimator, and the ML estimator and SNV estimator that rely on the pilot inserted in frame belong to data-aided (DA) estimator. In contrast, the M_2M_4 estimator is independent of the transmitted data and does not require carrier phase recovery and receiver decision.

978-1-7281-6992-7/20/$31.00 ©2020 IEEE

However, the conventional CSI estimation algorithms usually take a long time to obtain the current CSI and cannot sense the future trend. In this case, many articles have intensively researched on channel prediction, and proposed many classic channel prediction models based on machine learning. Yadan Zheng [13] proposed a modified ARIMA model for channel quality indication (CQI) prediction to solve the long delay problem in satellite environment. Furthermore, G. Liu [14] and J. Joo [15] used long short-term memory (LSTM) network to predict SNR in vehicle-to-everything (V2X) and V2V communication system. C. Luo [16] proposed an OCEAN model to predict CSI in 5G wireless communication system.

In high mobility scenario such as highway and HST, there emerge many challenges in predicting CSI. Specifically, the first problem is that the high speed movement of vehicles will cause rapid changes in channel characteristics. As J. Camp [10] mentioned, coherence time will decrease with the increasing speed, which easily causes fast fading in millimeter-wave system but the common regression algorithm such as ARIMA cannot track rapidly changing channel information. It usually leads to incorrect prediction results and makes the adaptive transmission scheme unable to select the correct modulation mode. Furthermore, the machine learning methods [14]–[16] mentioned above are all based on pilot sequences in the frame to extract CSI and the only way to adapt the fast changing channel is increasing the inserted pilots. However, the consequence is that the transmission efficiency is greatly reduced and the system complexity is increased. Therefore, we need to find an efficient way to perform CSI prediction in such high speed scenarios.

In this paper, we develop a novel blind channel information prediction model based on deep neural network (BCPMN) to adapt the fast changing channel and predict SNR without certain pilot. Unlike previous works that using pilot to estimate CSI then make prediction, our model only needs the raw receive signal as input instead of historical CSI or pilot data and the output is the future channel characteristics. To achieve our goal, we reshape the input data into a specific

2-dimensional matrix. The idea is to gather CSI and channel changing patterns at the same time, so that CNN can extract advanced features more easily. After reshaping and CNN processing, only the features related to channel changes are left, so LSTM can more accurately extract the time correlation of data at different times and predict CSI.

The paper is organized as follows. We first introduced the application scenario and the corresponding channel model in Section II. Then we describe the proposed prediction scheme in detail in Section III, including the data preprocessing method and the network architecture. Afterwards, we present the simulation results in Section IV, and finally conclude this paper in Section V.

II. System model

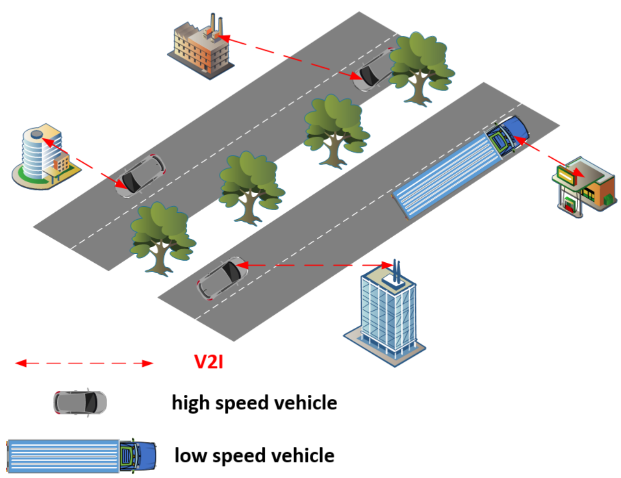


Fig. 1. A typical V2I communication scenario on the highway.

A typical V2I mm-wave communication scenario on the highway is shown in Fig.1, where the vehicles with different speeds communicate with base stations using mm-wave bands. In this scenario, the vehicles need to predict the rapid changes of CSI in real time, perform MEC and work with base stations to complete and the implementation of adaptive transmission scheme to improve the efficiency [8] and throughput [10] of wireless communication systems. To be more specific, we consider the SNR in the V2I channel as the prediction target than vehicles can use the predicted SNR to switch proper modulation modes to improve the communication quality.

Since there is no mature highway mm-wave channel model and the scenario is highly similar to HST environment. Therefore, in this work, we use a high speed trains wireless dynamic channel proposed by Y. Chang [4], which is based on a 60GHz mm-wave wireless system. In the HST scenario, the high speed trains are the vehicles and the base stations are placed on the beams directly above the rails. The envelope of the channel response is subject to the Rician distribution, and firstly the received signal model can be defined as:

$$r(k) = \sum_{l=0}^{L-1} h_l(k-l)s(k-l) + n(k), \tag{1}$$

where k represents the index of sample in time domain. $s(k)$ is the transmitted signal and $n(k)$ is the Gaussian noise. L is the number of multipath, and $h_l(k)$ is the lth multipath channel model. When $l = 0$, $h_0(k)$ is the LoS path as below:

$$h_0(k) = \sqrt{\frac{K}{K+1}} \frac{x^{\text{LoS}}}{|x^{\text{LoS}}|} \times e^{j2\pi f_d k T_s \cos(\theta^{\text{LoS}})\cos(\phi^{\text{LoS}})}, \tag{2}$$

$$f_d = \frac{v f_c}{c}, \tag{3}$$

where K is the Rician factor, x^{LoS} is a complex Gaussian variable with zero mean and unit variance, T_s is the sampling period, f_d represents the maximum Doppler shift, f_c is the carrier frequency, c is the light speed and v is the vehicle speed. Notation θ^{LoS} and ϕ^{LoS} represent the AoAs of the azimuth and elevation. When $0 < l < L$, $h_l(k)$ is the lth NLoS path:

$$h_l(k) = \sqrt{\frac{1}{K+1}} \frac{h'_l(k)}{\alpha}, \tag{4}$$

$$h'_l(k) = x^{\text{NLoS}} \times e^{j2\pi f_d k T_s \cos(\theta^{\text{NLoS}})\cos(\phi^{\text{NLoS}})}, \tag{5}$$

where x^{NLoS} is multipath channel gain and α is the power normalization parameter:

$$\alpha = \sqrt{\sum_{l=1}^{L-1} |h'_l(0)|^2}, \tag{6}$$

In the HST channel, our goal is to accurately predict the change of SNR with minimum frame overhead. From Eq. (3), we can see that in the millimeter wave system, due to the high carrier frequency, the maximum Doppler shift f_d will be greatly increased as the vehicle speed increases. As a consequence, the coherence time T_c will decrease, which is defined as:

$$T_c = \frac{0.423}{f_d}, \tag{7}$$

when the symbol period is greater than the coherence time T_c, fast fading occurs, which denotes that channel change rate is faster than symbol rate. However, the pilot-based channel estimation and prediction algorithms need to increase the insertion density of the pilot to adapt this rapidly changing channel, which will significantly reduce the system transmission efficiency and estimation accuracy in the rapidly changing channel. Therefore, we propose a blind channel prediction algorithm that can adapt to rapidly changing channels and does not rely on pilots, while the objective function can be expressed as:

$$G = \arg\min_{f,N,L} \|n - f(R_{N\times L})\|_2^2, \tag{8}$$

where n is the true SNR value, f represents the function that the neural network needs to fit, R is the received data matrix and $N \times L$ is the matrix dimension. $R_{N\times L}$ is the matrix obtained from the received signal sequence after data preprocessing.

III. Channel Prediction Scheme

In this section, we propose a BCPMN channel information prediction scheme based on CNN, LSTM and DNN, where only the received signal is needed to predict the future SNR. The conventional waveforms, like those in 802.11ad [17], use pilots to estimate transmission power and SNR. However, they cannot adapt to the rate of change of rapidly changing channel and the increased payload cost of pilot is unacceptable. For a certain modulation mode, its power statistics are regular, so our scheme can use neural network to learn signal power statistics from raw received signal then estimate SNR. In fading channel, the channel parameters change randomly, but within a relatively small time window, this change can be fitted and predicted, so the main idea of this scheme is to use CNN's excellent capability on feature extraction and LSTM's remarkable ability on time sequences analysis to extract the channel information hidden in the received signal. We first preprocess the data so that the hidden channel features can be better extracted by BCPMN. Then we fed the reshaped data into the BCPMN to predict the CSI we need.

A. Data Preprocessing

In this subsection, we propose a data preprocessing method to convert the raw data into the format we need. Since we need to extract the time correlation of channel information changes hidden in received signal, data sent to the network needs to be sufficient and chronological. Assuming that the channel state remains stable within the same frame. As is shown in Fig. 2. When the receiver detects a data frame, we randomly truncate and buffer a signal with a fixed time window length L. Then we stitch the cached data of N adjacent windows into a two-dimensional matrix $\widetilde{R}_{N\times L}$, which can defined as:

$$\begin{aligned} R_t &= [r_{t,1}, r_{t,2}, \cdots, r_{t,L}], \\ \widetilde{R}_t &= [R_{t-N+1}^{\mathrm{T}}, \cdots, R_{t-1}^{\mathrm{T}}, R_t^{\mathrm{T}}]^{\mathrm{T}} \end{aligned} \tag{9}$$

This matrix is the input data matrix we need, where the $R_{1\times L}$ is the truncated data of each frame, r is a sample, t is the index of frame and N is the time step. The value of N refers to the pilot length of 802.11ad [17] and L is determined by minimizing the validation loss in simulation experiments.

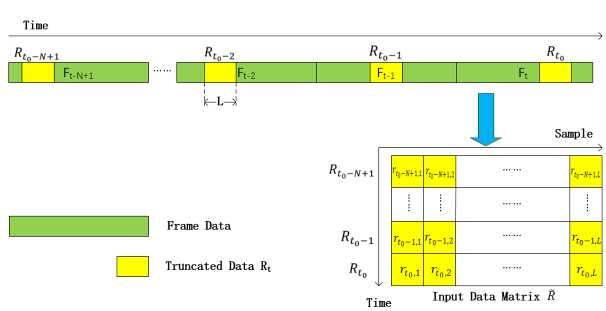


Fig. 2. The 1-dimensional data of t_0-th frame is converted into a 2-dimensional matrix in the data preprocessing scheme.

The main idea of constructing the matrix in this way is to make full use of CNN's feature extraction capabilities. The convolutional layer uses a convolution kernel to multiply its value element-by-element with the input matrix and then sum them up. Each row vector of the input matrix represents the channel information of different frames, and the change trend of channel information is implied between each row vector. Similar to the basic image edge detection problem [18], when the channel information of adjacent row vectors changes, the convolutional layer can extract the edge features between adjacent row vectors. The edge features represent the change trend and also the short-term time correlation of channel information. Therefore, when the input matrix is fed into CNN, the channel information of each frame and the short-term time correlation of adjacent frames can be well extracted. The rapid change of channel information makes the edge more obvious, and the extracted edge features are also clearer, so our model can achieve good performance in tracking fast changing channel.

B. Structure of the BCPMN

The structure of the BCPMN is shown in Fig. 3. The network consists of two 2-D convolutional layers, four LSTM layers and two fully connected layers. The motivation of using the proposed neural network structure is to make it easier for LSTM to fit and predict the hidden channel information from the raw received signal.

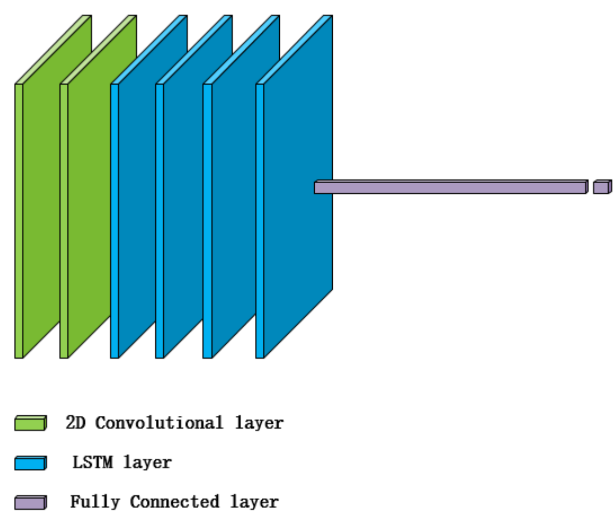


Fig. 3. The BCPMN model contains of 2 convolutional layers, 4 LSTM layers and 2 fully connected layers.

Firstly, the input data matrix $\widetilde{R}$ is sent to CNN. Similar to Yoon Kim [19] using CNN and LSTM for natural language processing, we consider a row vector of $\widetilde{R}$ as a word and all row vectors from R_{t-N+1} to R_t are treated as a sentence, which is a similar format with the input in [19]. Thus the first 2-D convolutional layer can extract the low-level features from the adjacent samples r and the adjacent frames R. The low-level features represent the channel information of each frame and are similar to the meaning of a word. Then the second convolutional layer extracts the change trend of channel information from the low-level features, which is similar to the edge and correlation feature extraction of adjacent words in a sentence. The convolutional operation can be described as Eq. (10), where $\widetilde{R}^k$ is the input matrix of the kth channel, F^{kl} is the convolution kernel in row k and column l and $\widetilde{Y}^l$ is the output matrix of the lth channel. Assuming that the convolutional layer has L output channels, K input channels and the kernel size is $I \times J$. Relu is nonlinear activation function and b is a bias and Y^l is the output of CNN. The irrelevant features in the input matrix have been removed, and the output features of CNN still retain time correlation.

$$\begin{aligned}\widetilde{Y}^l(m,n) &= \widetilde{R}^k(m,n) * F^{kl}(m,n)\\ &= \sum_{k=0}^{K-1}\sum_{i=0}^{I-1}\sum_{j=0}^{J-1}\widetilde{R}^k(m+i,n+j)F^{kl}(i,j),\\ Y^l &= \text{Relu}(\widetilde{Y}^l + b)\end{aligned} \tag{10}$$

Due to the fact that CNN's receptive field is limited by the size of the convolution kernel, it is impossible to capture the long-term dependence of channel information and difficult for CNN to accurately fit and predict it. Then the LSTM layers are used to exploit the long-term change trend in the whole time steps of $\widetilde{R}$. Assuming that y_t is the output from CNN and the input of LSTM in time t and h_t is the output in time t. LSTM's long-term memory comes from cell state C_t, and the short-term memory comes from h_t. These two parameters can both be transmitted on the time axis, and the specific update mechanism is as follows. Firstly, the cell state C_{t-1} of the previous LSTM cell enters the current LSTM cell and selectively forgets something through a forget gate, which can be defined as:

$$f_t = \sigma(W_f \cdot [h_{t-1}, y_t] + b_f) \tag{11}$$

The input of forget gate is concatenated by h_{t-1} and x_t. Secondly, the input gate is used to update the current cell state C_t by following steps:

$$\begin{aligned} i_t &= \sigma(W_i \cdot [h_{t-1}, y_t] + b_i),\\ \tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, y_t] + b_C),\\ C_t &= f_t * C_{t-1} + i_t * \tilde{C}_t \end{aligned} \tag{12}$$

Then, the output gate selectively outputs the current short-term memory according to the current cell state C_t:

$$\begin{aligned} o_t &= \sigma(W_o \cdot [h_{t-1}, y_t] + b_o),\\ h_t &= o_t * \tanh(C_t) \end{aligned} \tag{13}$$

Finally, two fully connected layers are used to map high-level features back to low-level features and output the final predicted value. The first long fully connected layer is used to adapt the output dimension of LSTM, and the second short fully connected layer is used to output prediction results.

IV. Simulation results

In this section, we verify the performance of the proposed scheme in predicting SNR. Moreover, we compare the performance of different methods and test the generalization ability of the proposed model. The channel model we use is the Rician channel in article [4] and the channel parameters are shown in TABLE I. In this channel, The envelopes of the SNR and K factor follow a Gaussian distribution, and the multipath delay follows an exponential distribution with a parameter of 0.5×10^{-9}. The R_t we use is truncated from a piece of data at the end of the frame, which is a completely random sequence.

TABLE I
Channel Parameters

Parameter	Value
Modulation mode	$\pi/2$-QPSK
Symbol rate (Msym/s)	1760
Carrier frequency (GHz)	62.5
Number of multipath	2
Vehicle speed (Km/h)	320
Maximum Doppler shift (kHz)	18.5

TABLE II
Training Parameter Configuration

Parameter	Value
Training environment	python3.6.10, tensorflow-gpu1.13.1
Initial learning rate	0.001
LSTM dropout	0.3
CNN kernel size	3×3
Optimizer	adam
Loss	mse
Time step N	30
Truncated length L (symbol)	3328

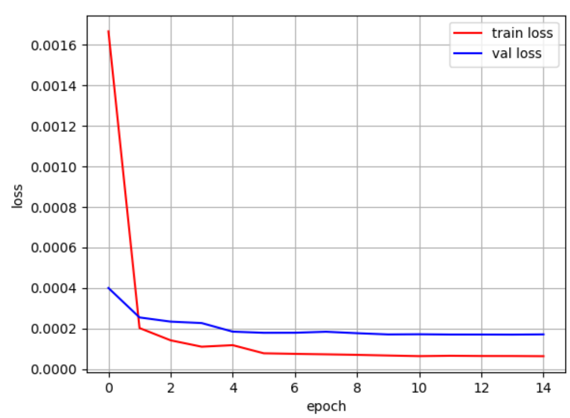
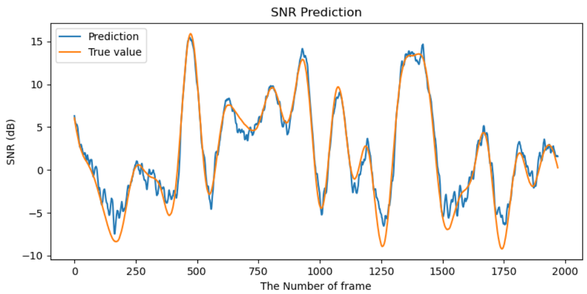
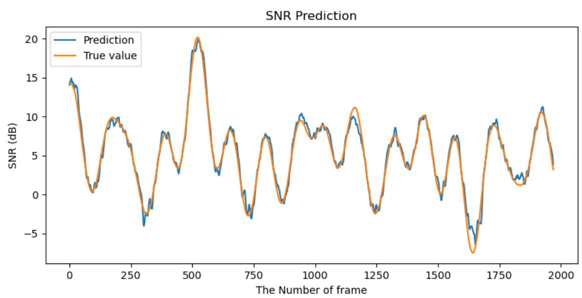


Fig. 4. The MSE loss curve of the BCPMN model.

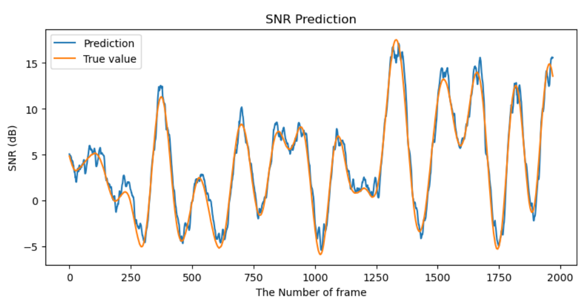
In order to improve the performance of the neural network and accelerate the convergence speed, we normalized the dataset into the range of $[0, 1]$. This paper uses python and tensorflow to build the neural network. The environment and main parameters of the neural network are shown in Table II.



(a) $v = 60km/h$, NMSE $= 0.039$



(b) $v = 120km/h$, NMSE $= 0.011$



(c) $v = 320km/h$, NMSE $= 0.020$

Fig. 5. SNR prediction results and NMSE at different vehicle speeds.

We use the normalized mean square error(NMSE) to calculate the error between the predicted value and the true value. The definition of NMSE is as follows:

$$\text{NMSE} = \frac{\sum_{i=1}^{n} |y_p(i) - y_t(i)|^2}{\sum_{i=1}^{n} |y_t(i)|^2} \tag{14}$$

where y_p is the predicted value and y_t is the true value. The dataset size is 20000 frames and the ratio of training set, verification set and test set is 0.8, 0.1, 0.1 respectively. Fig. 4 shows that the BCPMN can converge within 10 epochs during training.

Fig. 5 depicts the true and predicted SNR at different vehicle speeds. According to Eq. (3) and (7), the maximum Doppler shift in Fig. 5(a), Fig. 5(b) and Fig. 5(c) is 3472Hz, 6944Hz and 18519Hz, respectively. The channel change rate increases with speed. The average NMSE at each speed fluctuates slightly due to different test set. It can be seen from this figure that the predicted SNR is almost the same as the true value and can perform well at different speeds, which shows that the proposed model is very effective and can adapt well to different Doppler shift.

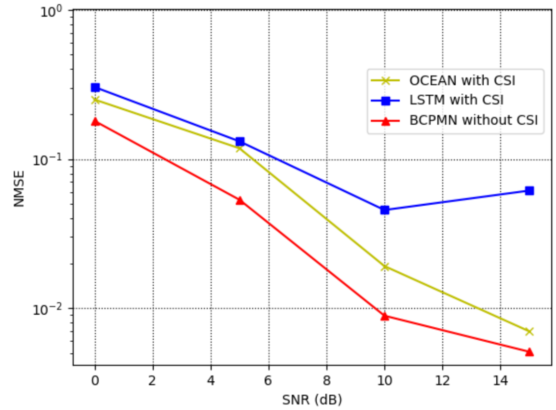


Fig. 6. The comparison of prediction performance between LSTM, OCEAN and BCPMN under different SNR.

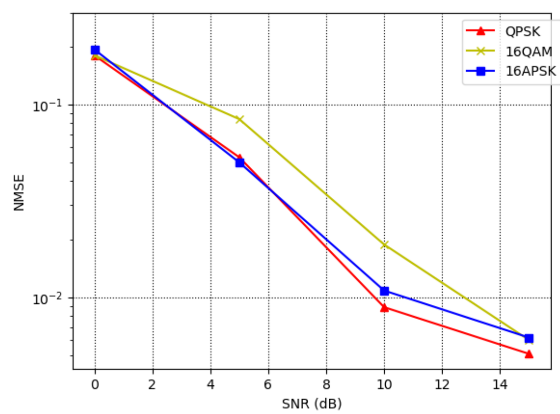


Fig. 7. Adaptability experiment of BCPMN model to QPSK, 16QAM and 16APSK modulation modes.

We compare the SNR prediction performance of our BCPMN scheme with the LSTM prediction scheme proposed by G. Liu [14] and the OCEAN scheme under the same simulation scenario in Fig. 5(c). Then we calculate the average NMSE and the NMSE under different SNR. In the simulation, both comparison algorithms use the estimated SNR based on the pilot as the input CSI. Fig. 6 shows the comparison results. The average NMSE of BCPMN, OCEAN and LSTM is 9.14×10^{-3}, 1.7577×10^{-2} and 5.9246×10^{-2} respectively. Because the signal features are more obvious under high SNR, the prediction error of each algorithm decreases with the increasing SNR. However, the prediction accuracy of the LSTM model is not improved when the SNR is high enough. It illustrates that the model cannot track highly dynamic changes. Simulation results show that the BCPMN model can achieve better performance than the LSTM model and the OCEAN model but our scheme only requires received signal instead of historical CSI.

Adaptive modulation and coding is widely used in the vehicular communication [10] and DVB-S2 system, including BPSK, QPSK, 8PSK, 16QAM, 16APSK and 32APSK modulation mode. Thus we select QPSK, 16QAM and 16APSK to test the generalization performance of our model. We have regenerated the dataset with the three different modulation modes under the same simulation scenario in Fig. 5(c). Then we have retrained the networks and tested the signals of the modulation modes. Fig. 7 shows the performance of our proposed model in QPSK, 16APSK and 16QAM modulation mode. The average NMSE of QPSK, 16QAM and 16APSK is 9.14×10^{-3}, 1.155×10^{-2} and 1.962×10^{-2} respectively, which shows that the BCPMN can achieve similar prediction performance in all three modulation modes and has good generalization performance.

V. Conclusion

In this paper, we proposed a blind channel information prediction model based on deep neural network in mm-wave wireless communication system. The BCPMN scheme can adapt to rapidly changing channel at different speeds and achieve accurate CSI prediction performance. We propose a data preprocessing method for the received signal, so that CNN can better extract the channel information. We compared the prediction performance of the three algorithms BCPMN, LSTM and OCEAN. The BCPMN can achieve a lower NMSE and only needs received signal instead of estimating the CSI in advance. To validate BCPMN's generalization performance, we conducted experiments in QPSK, 16QAM and 16APSK modulation mode and the results demonstrate that BCPMN can achieve similar prediction performance in all three modulation modes.

References

[1] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of Millimeter Wave Communications for Fifth-Generation (5G) Wireless Networks—With a Focus on Propagation Models," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6213–6230, 2017.

[2] C. F. Mecklenbrauker, A. F. Molisch, J. Karedal, F. Tufvesson, A. Paier, L. Bernado, T. Zemen, O. Klemp, and N. Czink, "Vehicular Channel Characterization and Its Implications for Wireless System Design and Performance," *Proceedings of the IEEE*, vol. 99, no. 7, pp. 1189–1212, 2011.

[3] J. Gozalvez, M. Sepulcre, and R. Bauza, "IEEE 802.11p vehicle to infrastructure communications in urban environments," *IEEE Communications Magazine*, vol. 50, no. 5, pp. 176–183, 2012.

[4] Y. Chang, M. Furukawa, H. Suzuki, and K. Fukawa, "Propagation Analysis with Ray Tracing Method for High Speed Trains Environment at 60 GHz," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, 2015, pp. 1–5.

[5] J. Kim and A. F. Molisch, "Enabling Gigabit services for IEEE 802.11ad-capable high-speed train networks," in *2013 IEEE Radio and Wireless Symposium*, 2013, pp. 145–147.

[6] Junhyeong Kim and Il Gyu Kim, "Distributed antenna system-based millimeter-wave mobile broadband communication system for high speed trains," in *2013 International Conference on ICT Convergence (ICTC)*, 2013, pp. 218–222.

[7] M. Agiwal, A. Roy, and N. Saxena, "Next Generation 5G Wireless Networks: A Comprehensive Survey," *IEEE Communications Surveys Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[8] L. Jiao, H. Yin, Y. Lyu, H. Huang, J. Dong, and D. Guo, "Optimal Schedule of Mobile Edge Computing Under Imperfect CSI," in *Algorithms and Architectures for Parallel Processing*, J. Vaidya and J. Li, Eds. Cham: Springer International Publishing, 2018, pp. 32–45.

[9] J. Zhao, Y. Liu, Y. Gong, C. Wang, and L. Fan, "A Dual-Link Soft Handover Scheme for C/U Plane Split Network in High-Speed Railway," *IEEE Access*, vol. 6, pp. 12 473–12 482, 2018.

[10] J. Camp and E. Knightly, "Modulation Rate Adaptation in Urban and Vehicular Environments: Cross-Layer Implementation and Experimental Evaluation," *IEEE/ACM Transactions on Networking*, vol. 18, no. 6, pp. 1949–1962, 2010.

[11] G. Holland, N. Vaidya, and P. Bahl, "A Rate-Adaptive MAC Protocol for Multi-Hop Wireless Networks," in *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '01. New York, NY, USA: Association for Computing Machinery, 2001, p. 236–251. [Online]. Available: https://doi.org/10.1145/381677.381700

[12] D. R. Pauluzzi and N. C. Beaulieu, "A comparison of SNR estimation techniques for the AWGN channel," *IEEE Transactions on Communications*, vol. 48, no. 10, pp. 1681–1691, 2000.

[13] Y. Zheng, S. Ren, X. Xu, Y. Si, M. Dong, and J. Wu, "A modified ARIMA model for CQI prediction in LTE-based mobile satellite communications," in *2012 IEEE International Conference on Information Science and Technology*, 2012, pp. 822–826.

[14] G. Liu, Y. Xu, Z. He, Y. Rao, J. Xia, and L. Fan, "Deep Learning-Based Channel Prediction for Edge Computing Networks Toward Intelligent Connected Vehicles," *IEEE Access*, vol. 7, pp. 114 487–114 495, 2019.

[15] J. Joo, M. C. Park, D. S. Han, and V. Pejovic, "Deep Learning-Based Channel Prediction in Realistic Vehicular Communications," *IEEE Access*, vol. 7, pp. 27 846–27 858, 2019.

[16] C. Luo, J. Ji, Q. Wang, X. Chen, and P. Li, "Channel state information prediction for 5g wireless communications: A deep learning approach," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 1, pp. 227–236, 2020.

[17] "IEEE Standard for Information technology–Telecommunications and information exchange between systems–Local and metropolitan area networks–Specific requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band," *IEEE Std 802.11ad-2012 (Amendment to IEEE Std 802.11-2012, as amended by IEEE Std 802.11ae-2012 and IEEE Std 802.11aa-2012)*, pp. 1–628, 2012.

[18] S. Xie and Z. Tu, "Holistically-nested edge detection," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1395–1403.

[19] Y. Kim, Y. Jernite, D. Sontag, and A. M. Rush, "Character-aware neural language models," in *Thirtieth AAAI conference on artificial intelligence*, 2016.