Spectrum Efficiency Prediction for Real-World 5G Networks Based on Drive Testing Data

Zheng Xing[1†], Haoyun Li[1†], Wenjie Liu[1], Zixiang Ren[1], Junting Chen[1], Jie Xu[1], and Cai Qin[2]
[1]School of Science and Engineering (SSE) and Guangdong Provincial Key Laboratory of Big Data Computing,
The Chinese University of Hong Kong, Shenzhen, 518172, China
[2]Huawei Technologies Co., Ltd, Shenzhen, 518129, China
Email: {zhengxing@link., haoyunli@link., juntingc@, xujie@}cuhk.edu.cn, qincai@huawei.com

***Abstract*—This paper studies the problem of predicting the spectrum efficiency (SE) for massive multiple-input multiple-output (MIMO) empowered 5G networks based on the reference signal received power (RSRP) collected from the drive test (DT). This problem is challenging because there is no precise model between the RSRP and the SE. The SE not only depends on the RSRP, which only captures the statistic of the channel, but also the beamforming strategy of the serving base station (BS) and the interference from the neighboring cells, which are not measured at the 5G client. This paper adopts a model-assisted data-driven approach to develop a machine learning model for the SE prediction. Specifically, a joint interference and SE prediction network is built, demonstrating prediction improvement over pure data-driven neural networks. In addition, a classification-assisted SE prediction network is constructed, which substantially reduces the prediction error at the low SE regime with marginally compromising the total prediction error. It is found that the model-assisted approach generally enhances the SE prediction accuracy by 2% approximately over a purely data-driven approach.**

I. Introduction

5G new radio (NR) wireless communication networks have provided orders of magnitude performance improvements with the densification of networks and massive multiple-input multiple-output (MIMO) deployment. Under massive MIMO, the base stations (BSs) operate sophisticated beamforming algorithms and resource allocation strategies to serve a massive amount of users. As a result, it becomes incredibly challenging for the operator to optimize the network parameters, such as tuning the tilting angles of the antenna panels at the BSs, for improving the spectrum efficiency (SE).

One promising strategy to optimize the 5G network under active research is to divide the network optimization into two parts: First, we develop a model to map the network parameters to the channel quality at each geo-location. Second, we build a model to predict the SE at each location based on the estimated channel quality. As a result, the operator can optimize the network parameters based on the predicted SE in the serving area. This paper focuses on the second part, and it aims at building a machine learning model to predict the SE based on the reference signal received power (RSRP).

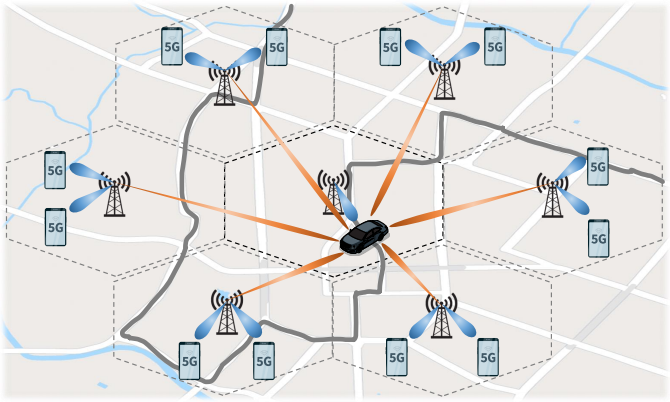

Figure 1: Illustration of the drive testing for 5G NR networks with densely deployed BSs and massive MIMO.

In the literature, there have been prior works investigating the prediction of long-term channel statistics (such as RSRP and channel state information (CSI) statistics) for real-world cellular networks. For instance, the authors in [1]–[3] studied the prediction of RSRP in LTE systems, indoor Wi-Fi networks, and vehicle networks, respectively. The studies in [4]–[6] considered the prediction of MIMO CSI by using partially measured information, and those in [7]–[9] considered the prediction of inter-cell interference. Besides, several attempts have been made to explore the relationship between SE and system parameters in massive MIMO networks (see, e.g., [10]–[13]). While these preliminary results modeled or predicted the SE in ideal scenarios, it is still a largely uncharted but challenging problem to develop a model to predict the SE based on the long-term channel statistics in 5G NR networks.

The SE prediction based on channel statistics faces a number of technical challenges. First, only partial data is available. In our data collection campaign, only the RSRPs of selected CSI beams measured at selected antennas are available due to practical constraints on the measurement device. Second, some essential information is missing. For example, it was too difficult to collect direct data to measure inter-cell interference. Besides, in practice, the datasets are built at different dates and locations statistically non-identical.

† Equal first authorship.

The work was supported in part by the National Key R&D Program of China with grant No. 2018YFB1800800, the National Natural Science Foundation of China under grants No. U2001208, 61871137, 62171398, and 92067202, the Science and Technology Program of Guangdong Province under grant No. 2021A0505030002 and 2019QN01X895, and the Shenzhen Fundamental Research Program under grant No. 20210318123512002.

This paper develops three techniques to tackle the above mentioned challenges and a series of machine learning models to demonstrate performance enhancement. First, we propose a location-based feature smoothing technique to ease the training of the neural network (NN). Based on this, we develop a joint interference and SE prediction network to enhance the SE prediction via estimating the interference as an intermediate step, bringing down the SE prediction error from 17% to 14.9%. Third, we build a classification-assisted SE prediction network, which reduces the prediction error at the low SE regime from 74% to 28%, slightly compromising the total SE prediction error.

II. System Model

A. SE Model for 5G MIMO Networks

Consider a cellular network with $Q+1$ BSs, where BS $q=0$ only serves a DT user and BS $q=1,2,\ldots,Q$ each serves multiple users as shown in Fig. 1. Each BS has N_{t} antennas, and the DT user has N_{r} antennas. The downlink channel from the qth BS to the DT user is given by $\mathbf{H}_q \in \mathbb{C}^{N_{\mathrm{r}}\times N_{\mathrm{t}}}$. Let $\widetilde{\mathbf{x}}_q \in \mathbb{C}^{d_q}$ denotes the message transmitted at the qth BS, $q=0,1,\ldots,Q$, where d_q are the numbers of data streams and $\mathbb{E}\{\widetilde{\mathbf{x}}_q\widetilde{\mathbf{x}}_q^{\mathrm{H}}\}=\mathbf{I}$. Let $\mathbf{V}_q \in \mathbb{C}^{N_{\mathrm{t}}\times d_q}$ denotes the transmit precoding matrix at the qth BS, where $\mathrm{tr}\{\mathbf{V}_q\mathbf{V}_q^{\mathrm{H}}\}\le P$ with P denoting the total power constraint and $\mathrm{tr}(\cdot)$ denoting the trace. The received signal at the DT user is given by

$$\mathbf{y}=\mathbf{H}_0\mathbf{V}_0\widetilde{\mathbf{x}}_0+\sum_{q=1}^{Q}\mathbf{H}_q\mathbf{V}_q\widetilde{\mathbf{x}}_q+\mathbf{n}$$

where $\mathbf{n}\sim\mathcal{CN}(0,\sigma^2\mathbf{I}_{N_{\mathrm{r}}})$ is the additive complex Gaussian noise.

While the transmission mechanism is complicated in an actual 5G network, we only focus on the first-order behavior of the transmission system and employ a simple de-correlation model to assist the design of the NN. Specifically, consider that the DT user employs a zero-forcing de-correlator and uses $\mathbf{u}_i$ to extract the ith data stream.[1] As a result, the received signal-to-interference-and-noise ratio (SINR) of the ith data stream is given by

$$\gamma_i=\frac{||\mathbf{u}_i^{\mathrm{H}}\mathbf{H}_0\mathbf{v}_{0,i}||^2}{\sigma^2+\sum_{q=1}^{Q}\|\mathbf{u}_i^{\mathrm{H}}\mathbf{H}_q\mathbf{V}_q\|^2} \tag{1}$$

where $\mathbf{v}_{0,i}$ is the ith column of the precoding matrix $\mathbf{V}_0$ for the DT user and $\|\cdot\|$ denotes the Euclidean norm. Thus, the instantaneous downlink SE for the DT user can be roughly modeled as

$$s(\{\mathbf{H}_q\})=\sum_{i=1}^{d_0}\log_2(1+\gamma_i) \tag{2}$$

where d_0 denotes the transmission rank or the number of data streams transmitted.

[1]In the low signal-to-noise ratio (SNR) regime, the BS may prefer to transmit only a single data stream, where zero-forcing can be identical to match filtering. At high SNR, zero-forcing is optimal.

The BS is believed to use the "best effort" to design the precoding matrix $\{\mathbf{V}_q\}$ based on the instantaneous information measured and reported by the DT user, such as RSRP, channel quality information (CQI), and rank indicator (RI), as well as the transmission history, *e.g.,* due to the implementation of the hybrid automatic repeat request (HARQ) scheme [14]. The goal of the paper is to predict the average downlink SE of the DT user $\bar{s}\triangleq\mathbb{E}\{s(\{\mathbf{H}_q\})\}$ based on the statistical information, RSRP, specified in the following subsection, where the expectation is taken over the small-scale fading of the MIMO channel at the order of hundreds of milliseconds.

B. Measurement Model for RSRP

Denote $\mathbf{Z}\in\mathbb{C}^{N_{\mathrm{t}}\times N_{\mathrm{t}}}$ as the discrete Fourier transform (DFT) matrix with its ith column $\mathbf{z}_i$ being the ith *CSI beam*. Consider to partition the index set $\{1,2,\ldots,N_{\mathrm{t}}\}$ into $M=N_{\mathrm{t}}/4$ subsets $\mathcal{C}_1,\mathcal{C}_2,\ldots,\mathcal{C}_M$.[2] Define $\tilde{\mathbf{z}}_i\triangleq\frac{1}{4}\sum_{j\in\mathcal{C}_i}\mathbf{z}_j$ as the beamforming vector for the synchronization signal (SS). It is clear that the beamwidth of the SS is wider than that of the *CSI beam*.

The CSI-RSRP measured at the jth receive antenna for the ith *CSI beam* transmitted by the qth BS is defined as the average received signal power $\bar{g}_{q,i}^{[j]}=\mathbb{E}\{||\mathbf{e}_j^{\mathrm{T}}\mathbf{H}_q\mathbf{z}_i||^2\}$, where the superscript T denotes the transpose, $\mathbf{e}_j$ is a vector of zeros except for the jth entry being 1, and the expectation $\mathbb{E}\{\cdot\}$ is taken over the small-scale fading. Similarly, the SS-RSRP measured at the jth receive antenna for the ith *CSI beam* transmitted by the qth BS is defined as $\tilde{g}_{q,i}^{[j]}=\mathbb{E}\{||\mathbf{e}_j^{\mathrm{T}}\mathbf{H}_q\tilde{\mathbf{z}}_i||^2\}$.

However, due to practical constraints, only partial information of the RSRPs is available. Let $\{\bar{g}_{q,(i)}^{[j]}\}$ be the *ordered* CSI-RSRP, such that $\bar{g}_{q,(1)}^{[j]}\ge\bar{g}_{q,(2)}^{[j]}\ge\cdots$, and $\{\tilde{g}_{q,(i)}^{[j]}\}$ be the ordered SS-RSRP. For the CSI-RSRP measurement related to the serving cell $q=0$, only the $N_s=8$ strongest CSI-RSRPs $\bar{g}_{q,(1)}^{[j]},\bar{g}_{q,(2)}^{[j]},\ldots,\bar{g}_{q,(N_s)}^{[j]}$ are available. For the RSRP measurement related to the neighbor cells, only the strongest SS-RSRP $\tilde{g}_{q,(1)}^{[j]}$ is recoded, which is denoted as $\tilde{g}_{q,(1)}^{\max}$, and the neighboured $Q=6$ cells are considered. Furthermore, only two receive antennas $j=1,2$ record the CSI-RSRP. So, our problem is to predict SE based on the RSRP information $\{\bar{g}_{0,(i)}^{[1]},\bar{g}_{0,(i)}^{[2]}\}_{i=1}^{N_s},\{\tilde{g}_{q,(1)}^{\max}\}_{q=1}^{Q}$.

C. Data Collection

We collect the RSRPs of the CSI beams of the serving cell and the RSRPs of the SS from neighbor cells by using a 5G phone. An auxiliary equipment is used to measure the SE, and the SE is defined as the instantaneous media access control (MAC) layer throughput over one second divided by the number of the resource blocks (RBs) used.

In addition, to assist the development and training of our NN model, a massive amount of auxiliary data is collected by using the equipment mentioned above, including the date, time, and GPS location of the data sample collected, average RI, downlink initial block error rate (BLER), and average modulation and coding scheme (MCS). The data density typically ranges from 1 to 5 samples per second, depending on the attributes

[2]In practice, N_{t} is a multiple of 4.

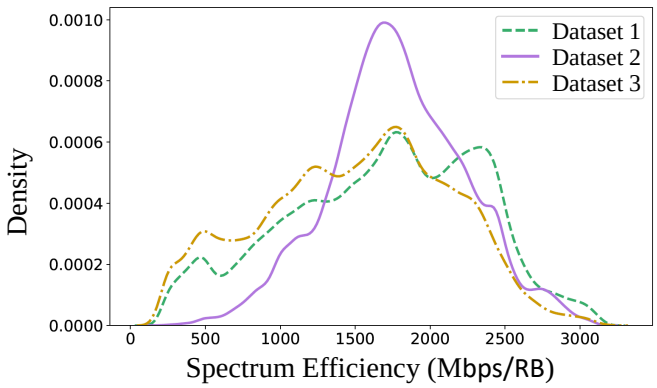

Figure 2: The SE distribution of the three datasets. The statistics are non-identical.

of the data. For example, RI and MCS are recorded once per second, and RSRPs are updated 3-5 times per second.

The data collection campaign was conducted in both Chengdu and Shenzhen, two major cities in China. Three datasets are formed:

- Dataset 1: Collected from April to October 2020 in Chengdu with 87-hour duration and 142 km travel distance;
- Dataset 2: Collected in December 2020 in Chengdu with 19-hour duration and 27 km travel distance, where the network was operated under different antenna configurations from Dataset 1;
- Dataset 3: Collected from August to September 2020 in Shenzhen with 12-hour duration and 19 km travel distance;

Statistics analysis shown in Fig. 2 reveals two challenges: First, the statistics are quite different across datasets, which makes it difficult to generalize the NN model for the SE prediction to a new wireless network (e.g., Dataset 2) at a new geographical location (e.g., Dataset 3). Second, the distribution of the SE is highly uneven, and there is little data in the low SE regime, which makes it challenging to make an accurate prediction when the SE is small. We will address these challenges in Sections IV and V, respectively.

III. Design of a Preliminary NN Model

A. Data Preprocessing

The data collected is noisy and contains a large amount of missing entries due to two factors: First, the timestamps of different features reported by the service are misaligned. For example, there can be a 500-millisecond lag between the SE data and the nearest RSRP or CQI record. Second, the DT user that collects the data is a vehicle moving at variable speed in real traffic. In addition, there is a significant amount of "outliers" due to transmitting small packets for hand-shaking from higher layers.

Based on these observations, we perform data interpolation, outlier removal, and normalization as follows.

- **Interpolation**: As different data entries are recorded asynchronously, we perform linear interpolation with respect to time to fill in the missing values of each feature.
- **Outlier removal**: We remove data with strong channel quality but low SE or initial BLER. Such a sample is probably due to small packet transmission from the higher layer, but this does not represent the achievable SE of the system. Specifically, we remove data that has CSI-RSRP above -60 dBm, SE lower than 100Mbps/RB or higher than 3200Mbps/RB, and initial BLER lower than 5%.
- **Normalization**: Each feature is normalized to the range [0, 1] to speed up NN learning and ensure that each feature contributes approximately proportionately to the NN. Specifically, denote M_j and m_j, respectively, as the maximum and minimum values of the jth feature in the training set. Then, a data sample x of feature j is normalized as $\bar{x} = \frac{x-m_j}{M_j-m_j}$.

B. Location-based Feature Smoothing for Training

To further prevent the NN over-fitting to the noise, we propose to smooth the *training* data. Note that user performance in a cellular network is significantly correlated with the location since the signal propagation depends on the local environment. Therefore, the feature smoothing scheme should exploit the SE performance's time and spatial correlation. With such a goal, we design a 1D filter as follows.

Let o_i be the cumulated travel distance of the DT user when the ith data sample $\mathbf{x}_i$ is collected, where $\mathbf{x}_i$ is the interpolated feature vector. Note that the temporal adjacency is also captured in the sequence o_i, since the vehicle speed is bounded. For a data sample $\mathbf{x}$ collected at the cumulated travel distance o, the smoothed version $\hat{\mathbf{x}}$ is computed based on the weighted sum over a subset of sample indices $\Omega = \{k : |o - o_k| \leq \Delta\}$ collected in its Δ-neighborhood:

$$\hat{\mathbf{x}} = \frac{\sum_{k\in\Omega} w_k \mathbf{x}_k}{\sum_{k\in\Omega} w_k}$$

where the weights $w_k = e^{-\frac{|o-o_k|}{\Delta/2}}$ are determined by the difference of the travel distance. It is clear that a high weight implies closeness in both time and space.

It is found that the location-based feature smoothing for the training data does improve the SE prediction performance in the test set, as seen in Table I.

C. A Preliminary NN Model

We adopt a multi-layer perceptron (MLP) NN to serve as a benchmark to build a preliminary NN, namely SENet, shown in Fig. 3. As mentioned in Section II-B, the input of the model is 22 RSRPs $\{\bar{g}^{[1]}_{0,(i)}, \bar{g}^{[2]}_{0,(i)}\}_{i=1}^{8}, \{\tilde{g}^{\max}_{q,(1)}\}_{q=1}^{6}$. There are three hidden layers, each consisting of 32 nodes with the rectified linear unit (ReLU) as the activation function.

The loss function is designed as a *density-weighted relative loss* to address the unevenness of the training data. Specifically, the loss function is defined as:

$$L_{\text{SENet}} = \sum_i w_i \left|\frac{\widetilde{y}_i - y_i}{y_i}\right|$$

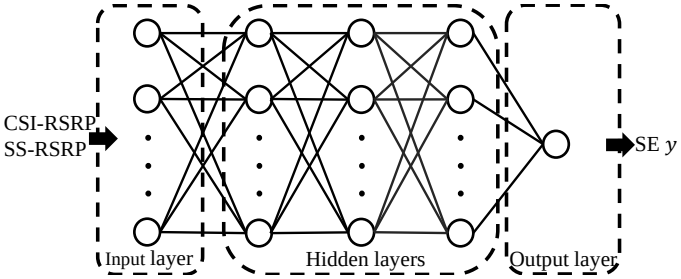

Figure 3: Architecture of the SENet.

where $\widetilde{y_i}$ is the predicted SE of the ith sample, and y_i is the SE of the ith sample obtained in Section II-C.

The weights for each sample are adjusted according to the density of the SE, $w_i = \frac{1}{\hat{f}(y_i;h,\{y_i\}_{i=1}^n)}$, where $\hat{f}(y_i;h,\{y_i\}_{i=1}^n)$ is the empirical density function in the training set. Here, we adopt the kernel density estimator to construct the density function as $\hat{f}(y;h,\{y_i\}_{i=1}^n) := \frac{1}{nh}\sum_{i=1}^n K(\frac{y-y_i}{h})$, where $K(\cdot)$ is the Gaussian kernel function with smoothing parameter h and n is the total number of the sample.

The parameters of the model are set in Section VI. The performance is evaluated using mean absolute percentage error (MAPE) $e = \mathbb{E}\{|\tilde{y} - y|/|y|\}$, and we are interested in the accuracy on different SE regimes, i.e., the expectation $\mathbb{E}\{\cdot\}$ is conditioned on different SE criteria.

Table I: MAPE[%] of the SE prediction on the test sets for the SENet under different values of W.

W[meter]	25	30	35	40	45	50	55	60
Dataset 2	19.0	18.2	18.0	17.0	17.6	18.3	19.2	20.2
Dataset 3	33.6	32.4	31.2	30.7	32.2	33.3	34.2	35.1

Table II: MAPE[%] of the SE prediction on the test sets for the SENet with $W = 40$ under different SE regimes.

SE regimes	Low	Medium-low	Medium	High	Total
Dataset 2	71.0	18.0	12.0	13.0	17.0
Dataset 3	72.0	24.0	28.0	32.0	30.7

Table I summarizes the MAPE of the SE prediction for the MLP model under different values of W for feature smoothing. First, feature smoothing on the training data reduces prediction errors, and $W = 40$ meters yields the best performance on both test sets. Second, although the model is trained using the density-weighted relative loss to mitigate the unbalance of the training data, the MLP model still suffers inferior performance in the low SE regime. As shown in Table II, we divide the whole SE region into four regimes, namely low, medium-low, medium, and high regimes, by three partition indices 1000, 1500, and 2500. The MAPEs of the low SE regime of Dataset 2 and Dataset 3 are 71% and 72% separately, larger than the other SE regimes' error. In the remaining part of the paper, we study two techniques to improve both the overall accuracy and the accuracy of the low SE regime.

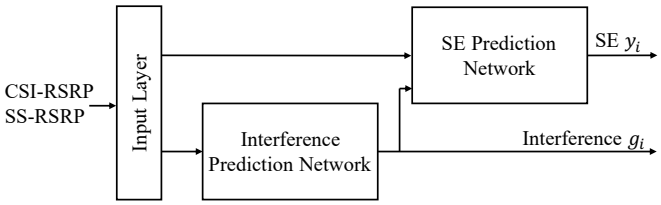

Figure 4: Architecture of the ISENet.

IV. Interference Model Aided SE Prediction

It is well-known that an MLP network with more than two hidden layers and sufficient neurons with non-linear activations can approximate any function. This section aims to improve the NN structure by exploiting the mathematical structure of the transmission model.

A. Joint Interference and SE Prediction

As shown in the transmission model for the 5G MIMO networks in Section II (see. (1) and (2)), the downlink SE depends on a number of factors, including the precoding matrices $\mathbf{V}_q$, the transmission rank d_0, and the SINR γ_i. Although there is a clear relationship among the *interference*, SINR, and the SE as in (1) and (2), we cannot obtain or predict the precoding matrices, and the transmission rank as instantaneous channel information is not available. By contrast, it is relatively easier to predict the average interference. Specifically, one may use the SS-RSRP from adjacent cells as an essential feature to predict the average interference level, which may help predict SE.

Towards this end, there are still two difficulties to be addressed. First, there is no auxiliary data that directly captures the interference in our dataset. As a result, we need to define a *proxy* that can be easily computed from the auxiliary data for the interference level. Second, even with knowledge of the interference level, we need to train a model that maps the interference level to the SE.

We build a joint interference and SE prediction network to tackle the above two challenges, namely ISENet, as shown in Fig. 4. In particular, the ISENet combines the SENet with a new interference prediction network that aims to predict the total interference (plus background noise), denoted by g that corresponds to the sum of $\sigma^2 + \sum_{q=1}^{Q} \|\mathbf{u}_i^{\mathrm{H}}\mathbf{H}_q\mathbf{V}_q\|^2$ in (1) from all the interference beams. The interference prediction network uses the CSI-RSRP and SS-RSRP as input, and the manually calculated interference value g is adopted as the training label, where g is calculated based on MCS and CSI-RSRP. As mentioned above, MCS comes from a parameter table that is preset by the BS manufacturer, and CSI-RSRP is measured by the 5G phone. The predicted interference value is then taken together with RSRP as input into the SE prediction network to estimate the SE. More specifically, for both the interference prediction network and the SE prediction network, there are three fully connected (FC) layers. Each activation function of the three FC layers is the ReLU function, and the

Table III: The MAPE[%] results before and after adding the low SE classification feature under Dataset 2.

	Low	M-low	Medium	High	Total
ISENet	74.0	22.0	13.0	14.0	14.9
cISENet	10.8	10.2	9.9	9.7	10.2

activation function of the output layer is the linear function. Moreover, the loss function of ISENet is

$$L_{\text{ISENet}} = \alpha \sum_i w_i |\frac{\widetilde{y}_i - y_i}{y_i}| + (1-\alpha) \sum_i w_i |\frac{\widetilde{g}_i - g_i}{g_i}|$$

where $\widetilde{g}_i$ denotes the predicted interference of the ith sample, g_i denotes the interference of the ith sample which is obtained in Section IV-B, and α is the proportional control factor. Here, $0 \leq \alpha < 0.5$ means that the loss of interference is more important than that of SE, and $0.5 < \alpha \leq 1$ means the opposite.

B. Generating the Proxy Interference Data

In realistic 5G NR networks, the actual interference value caused by neighboring cells is difficult to obtain, thus we use a mathematical model to manually get the estimated interference value as follows to tackle this issue. In particular, based on the SINR formula in (1), the interference g_i of the ith timestamp for DT is expressed as

$$g_i = \frac{P_{\text{r}}^{(i)}}{\widetilde{\gamma}_i}$$

where $P_{\text{r}}^{(i)}$ denotes the total received signal power of the ith sample and $\widetilde{\gamma}_i$ denotes the SINR of the ith timestamp for DT, which is different from the received SINR of the ith data stream γ_i. We approximate $P_{\text{r}}^{(i)}$ according to the CSI-RSRP from the primary cell as

$$P_{\text{r}}^{(i)} = \frac{1}{\eta^2} \sum_{j=1}^{\eta} \frac{1}{2} (\bar{g}_{0,(j)}^{[1]} + \bar{g}_{0,(j)}^{[2]}) \tag{3}$$

where η denotes the RI measured by the DT equipment that indicates the number of independent transmission layers d_0 used at the transmitter, or the number of independent channels between the transmitter and receivers. Notice that in (3), we choose the η strongest CSI-RSRP beam values, as it is highly likely the BS may choose to transmit the η data streams in the subspace spanned by the strongest η beams. Furthermore, we divide the average power of the η strongest beams by an additional factor of η, in order to capture the potential power allocation among the η transmission layers.

Next, we obtain the SINR value $\widetilde{\gamma}_i$ from the collected MCS based on the mapping table between the SINR and MCS values. Then convert $P_{\text{r}}^{(i)}$ in dBm, and the mathematical model of g_i is given by

$$g_i[dBm] = P_{\text{r}}^{(i)}[dBm] - \widetilde{\gamma}_i[dB].$$

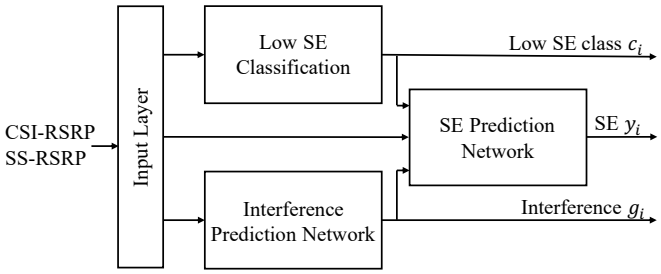

Figure 5: Architecture of the CISENet.

Therefore, the interference value is successfully estimated for the implementation of ISENet.

V. Classification-aided SE Prediction

The MAPE in the low SE regime can be reduced by adding a low SE classification prediction NN into the ISENet. The samples in the low SE regime are limited and noisy, which inspires the ISENet to pay more attention to the low SE regime. Notice that the ISENet model is used as a baseline model, and the classification-aided SE prediction NN, namely cISENet, is further designed to seek a better design structure for SE prediction model. Therefore, in the cISENet we design the NN same as that of ISENet, but the low SE classification feature is also utilized as a partial input of the total NN.

As shown in Table III, the MAPE of low SE goes from 74.0% to 10.8% after adding the low SE classification feature as the input of the ISENet, which is a significant improvement in our application.

Motivated by this, we present the new model as follows. First, we adopt the same structure as the SENet except for the output layer to design the low SE classification NN. There are two nodes in the output layer corresponding to two classes with softmax as the activation function. The cross-entropy loss function, which is calculated between the predicted category probability output and the actual category, is designed as follows:

$$L_{\text{C}} = -\sum_i w_i [c_i \times \log(p_i) + (1-c_i) \times \log(1-p_i)]$$

where c_i is the low SE classification label of the ith sample. If the ith sample is in the low SE regime, $c_i = 1$, otherwise $c_i = 0$. p_i and $1 - p_i$ are the output of the SE classification NN.

As shown in Fig. 5, the SE prediction model, namely CISENet, takes the output of the low SE classification NN c_i and the predicted interference g_i together with RSRP as input to estimate SE. Combining the ISENet's loss and the low SE classification loss, the loss function of the CISENet is set as

$$L_{\text{CISENet}} = L_{\text{ISENet}} + L_{\text{C}}.$$

VI. Experiment Results

This section presents experimental results to evaluate the performance of our proposed the SENet, ISENet, and CISENet, by using the three datasets mentioned in Section II-C. In the

Table IV: The comparison of the three proposed models.

	SE regime	Low SE		Medium-low SE		Medium SE		High SE		Total	
	SE range[Mbps/RB]	*0-1000*		*1000-1500*		*1500-2500*		*2500-3200*		*0-3200*	
	MAPE[%] or SDAPE[%]	MAPE	SDAPE	MAPE	SDAPE	MAPE	SDAPE	MAPE	SDAPE	MAPE	SDAPE
Dataset 2	SENet	71.0	15.9	18.0	14.2	12.0	15.1	13.0	13.1	**17.0**	15.0
	ISENet	**74.0**	14.1	22.0	13.9	13.0	14.8	13.6	12.2	**14.9**	16.7
	CISENet	**28.0**	15.8	25.0	14.4	17.0	15.6	16.0	13.0	**19.3**	16.3
Dataset 3	SENet	72.0	21.2	24.0	19.3	28.0	18.4	32.0	18.8	30.7	20.9
	ISENet	78.0	21.9	26.0	19.8	20.0	18.6	19.0	19.1	28.8	21.4
	CISENet	32.0	21.4	31.0	19.5	36.0	18.4	38.0	18.9	31.1	21.0

experiment, Dataset 2 (after system parameters reconfigured) is used as the test dataset to evaluate the robustness of the models if we change the antenna configurations. Dataset 3 (collected from a different city scheme) is used to evaluate the portability performance of the models under different areas. The model is trained using an Adam optimizer with an initial learning rate $\eta_0 = 0.001$, batch size 64, and fixed decay rate 0.9. The proportional control factor $\alpha = 0.8$, and the number of training epochs is 100. Our models are initialized with randomized weights.

Table IV shows the performance of SE prediction in terms of MAPE e and standard deviation of the absolute percentage error (SDAPE) $\sigma_e = \sqrt{\mathbb{E}\{(|\tilde{y}-y|/|y|-e)^2\}}$, which are obtained by averaging over 50 tests. First, consider the case with Dataset 2. It is observed that the ISENet achieves the MAPE of 14.9%, which is 2.1% lower than that by the SENet. This shows that by exploiting the predicted interference based on the mathematical structure of wireless transmission models, the ISENet driven by both data and model has a stronger generalization ability and thus leads to better SE prediction capability than the pure data-driven SENet. It is also observed that while the SENet and ISENet outperform the CISENet in terms of the total prediction errors, the CISENet leads to a much lower prediction error (MAPE of 28%) in the low SE regime than that of 71% in SENet and 74% in ISENet. This indicates the value of low SE classification NN in reducing the prediction error at the SE regime with slightly compromised overall SE prediction performance. Second, it is observed that the performance of the three models based on Dataset 3 is worse than that based on Dataset 2, which motivates us to do more research in the future to improve the portability performance of our models.

VII. Conclusion

In this paper, we adopted a model-assisted data-driven approach to building a deep neural networks (DNN) model to predict the SE in a 5G NR network based on 5G phone measurement statistic RSRP and the auxiliary equipment measurement statistics, including MCS, SE, RI, and so on. The challenge to address is that the SE depends on a lot of hidden factors that are not directly captured by the CSI data available. To circumvent the difficulty, we developed a location-weighted averaging approach to de-noise the measurement data and then designed a DNN architecture trained to predict the interference with noise and the SE jointly. It is found that the ultimate prediction accuracy on SE can be improved by using the predicted interference as the NN input. This is substantiated by the massive drive testing data collected in a real-world 5G NR network at Chengdu, China. The training and testing data are collected at different periods with different network parameter configurations. An overall prediction error of 14.9% on the SE is achieved.

References

[1] J. H. Seong and D. H. Seo, "Selective unsupervised learning-based Wi-Fi fingerprint system using autoencoder and GAN," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 1898–1909, Mar. 2020.

[2] J. Thrane, B. Sliwa, C. Wietfeld, and H. L. Christiansen, "Deep learning-based signal strength prediction using geographical images and expert knowledge," in *Proc. IEEE Global Commun. Conf.*, Dec. 2020, pp. 1–6.

[3] T. Nishio and H. Okamoto, "Proactive received power prediction using machine learning and depth images for mmWave networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2413–2427, Nov. 2019.

[4] W. Peng and W. Li, "Downlink channel prediction for time-varying FDD massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 1090–1102, Sep. 2019.

[5] Y. Yang, F. Gao, Z. Zhong, B. Ai, and A. Alkhateeb, "Deep transfer learning-based downlink channel prediction for FDD massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7485–7497, Dec. 2020.

[6] P. Dong, H. Zhang, and G. Y. Li, "Machine learning prediction based CSI acquisition for FDD massive MIMO downlink," in *Proc. IEEE Global Commun. Conf.*, Dec. 2018, pp. 1–6.

[7] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 192–201, Feb. 2013.

[8] T. Ren and R. La, "Downlink beamforming algorithms with inter-cell interference in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 10, pp. 2814–2823, Oct. 2006.

[9] B. Soret, A. De Domenico, S. Bazzi, N. H. Mahmood, and K. I. Pedersen, "Interference coordination for 5G new radio," *IEEE Wireless Commun.*, vol. 25, no. 3, pp. 131–137, Nov 2017.

[10] S. Han and I. Chih-Lin, "Achieving high spectrum efficiency on high speed train for 5G new radio and beyond," *IEEE Trans. Wireless Commun.*, vol. 26, no. 5, pp. 62–69, Oct. 2019.

[11] T. Van Chien, E. Bjornson, and E. G. Larsson, "Sum spectral efficiency maximization in massive MIMO systems: Benefits from deep learning," in *Proc. IEEE Int. Conf. Commun.*, Jul 2019, pp. 1–6.

[12] Y. Zhong, T. Q. Quek, and W. Zhang, "Complementary networking for C-RAN: Spectrum efficiency, delay and system cost," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4639–4653, May 2017.

[13] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Aug 2002.

[14] A. Ahmed, A. Al-Dweik, Y. Iraqi, H. Mukhtar, M. Naeem, and E. Hossain, "Hybrid automatic repeat request (HARQ) in wireless communications systems and standards: A contemporary survey," *IEEE Commun. Surv. Tutor.*, vol. 23, no. 4, pp. 2711–2752, Jul 2021.