Accepted for Presentation at 2022 IEEE International Symposium on Antennas and Propagation, Denver, CO, July 2022

Optimization of 5G Infrastructure Deployment Through Machine Learning

Ziheng Fu(1), Swagato Mukherjee(2), Michael T. Lanagan(1), Prasenjit Mitra(1,3), Tarun Chawla(2), and Ram M. Narayanan(1)

(1) The Pennsylvania State University, University Park, PA 16802, USA {zqf5070, mxl46, pum10, rmn12}@psu.edu

(2) Remcom Inc., State College, PA 16801, USA {swagato.mukherjee, tarun.chawla}@remcom.com

(3) L3S Research Center, 30167 Hannover, Germany

Abstract—**The application of machine learning for optimal deployment of 5G infrastructure, such as the position and the orientation of the antenna that help achieve the best signal coverage, is investigated in this paper. This avoids the need to perform on-site measurements or extensive software simulations. Multivariate Regression (MR) and Neural Network (NN) models were applied to predict the signal coverage in an indoor environment. The results showed that the average prediction error using NN for the case investigated is 7 dB for a 60-GHz operating frequency, whereas the error using the MR technique is lower than 6 dB. The unique aspect in our work is the integration of the clustering algorithm and the NN machine learning model for predicting indoor signal coverage.**

I. Introduction

With the large-scale commercial usage of 5G communications being the goal for the next decade, the reception of 5G signal is becoming an issue. This is due to the unique characteristics of the 5G signals whose frequency can be above 40 GHz in order to improve the communications capacity. However, these high signal frequencies undergo significant attenuation and are distorted while interacting with various materials in the propagation medium. This factor makes it more challenging to characterize the received signal pattern. Some work has been undertaken to predict the signal coverage for indoor and outdoor environments [1], [2] and to develop numerical models for the signal propagation of 5G networks [3]. This paper investigates machine learning approaches for the accurate and efficient prediction of signal coverage at a frequency of 60 GHz in an indoor environment and thus, to provide a strategy for the deployment of 5G infrastructure. Results are compared with traditional approaches. The representative scene investigated here is an apartment of floor area around 60 m^2, as shown in Fig. 1.

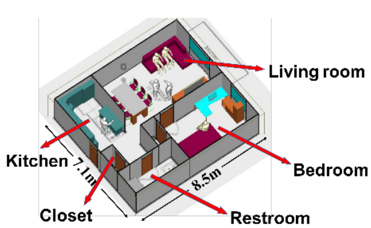

Fig. 1. View of the indoor environment, wherein the receivers and transmitters are evenly distributed within the rooms.

The apartment is assumed to contain 306 receivers evenly distributed within the room to measure the signal coverage and 306 transmitters to test the coverage corresponding to all possible deployments. One receiver and one transmitter are collocated as an integrated pair and the spacing between each transmitter-receiver pair is 0.5 m. The transmitter-receiver pair is mounted at a height of 1 m. The time cost will be significantly enormous if running the simulation for every single possible deployment of the transmitter is required; therefore, the machine learning algorithm is proposed.

II. Proposed Approaches

The merit of the signal coverage is defined as the weighted sum of the received signal strength (RSS) at all positions in the room. We assume that locations such as the living room and bedrooms require stronger signals than restrooms and the kitchen because the placement is supposed to be optimized for rooms where the user may be most often. Consider the m^{th} transmitter at location (x_t^m, y_t^m) is enabled and as a result the RSS of the n^{th} receiver at location (x_r^n, y_r^n) is $P_{(m,n)}$, the optimal placement $\left(\widehat{x_t^m}, \widehat{y_t^m}\right)$ is defined as:

$$\left(\widehat{x_t^m}, \widehat{y_t^m}\right) = \arg\max_m \sum_{n=1}^{N} w_{(x_r^n, y_r^n)} P_{(m,n)}, \tag{1}$$

where $w_{(x_r^n, y_r^n)}$ is the weight for the n^{th} receiver. The ground truth of received signal power is obtained from the Wireless InSite® ray-tracing simulation software which can provide accurate estimates of signal power in situ. The data collected are randomly divided into a training set and a testing set to verify the accuracy of the algorithms.

A. Neural Networks (NN)

1) Description

Neural networks are machine learning models widely used for nonlinear regression. The prediction of signal coverage is generalized so as to extract the relationship between the RSS and the position of the m^{th} transmitter and the n^{th} receiver. A neural network, shown in Fig. 2, was designed with three layers, including one hidden layer with a specific number of hidden units. The prediction of the signal power $\hat{P}$ can be described as

$$\hat{P}_{(m,n)} = W_2\left\{\text{tansig}\left[W_1 X_{(m,n)} + b_1\right]\right\} + b_2, \tag{2}$$

where W is the weight, b is the bias, $X_{(m,n)}$ is the normalized input space containing the coordinates of the transmitter and the receiver, and tansig[·] is the activation function. The coefficient in the network is updated to minimize the sum of the squares of the errors between the prediction of the signal power and the ground truth, obtained from Wireless Insite, using the Levenberg-Marquardt (LM) algorithm [4].

2) *K-Means Clustering*

The success of the design of neural networks may depend on reasonable values of initial weights and number of hidden units. While the attenuation of the wave along the propagation paths is inversely proportional to the square of the length of the path, the pattern of the received signal power will not strictly follow this pattern due to the interactions of high-frequency waves with walls and furniture. Consequently, it is appropriate to develop neural networks applied for different groups of receivers which are classified by K-Means Clustering based on the distance of propagation and the signal power collected from the simulation, rather than to design a single neural network to be applied for all the receivers. Since the number of unknown variables in the network is large, clustering facilitates network convergence. The number of clusters determined by the well-known elbow method is eight (8). The number of hidden units in each cluster is optimized through a number of trials.

B. *Multivariate Regression (MR)*

The multivariate regression is to obtain the relationship between independent variables by minimizing the sum of squares of error of prediction. The model of regression is expressed as

$$Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k, \qquad (3)$$

where Y_i denotes the output which is a function of inputs x_k, and the features of the input space is the same as that in Neural Networks. The β_ks are parameters to be estimated by ordinary least square regression, using

$$\widehat{\boldsymbol{\beta}} = (X^T X)^{-1} X^T y. \qquad (4)$$

III. Results

A. *Performance of MR and NN*

The performance of the MR and the NN models are compared by plotting the root-mean-square error (RMSE) of the prediction as a function of the size of the training set, as shown in Fig. 3. The testing error using the NN has an average RMSE of 7 dB while the MR gives 5.9 dB. The fluctuating plot of the error for the NN method might come from the randomness of the initial weights and the fact that the NN depends on the appropriate choice of the number of hidden units.

B. *Prediction of Signal Coverage*

The signal coverage corresponding to the transmitter placed at a particular position in each room can be predicted by either method discussed in the previous sections. The best position to locate the transmitter could then be determined by the average of predicted signal power, as shown in Fig. 4 for the entire apartment. The yellow pixels indicate that transmitters placed in those areas will provide better signal coverage compared to transmitters deployed in regions with dark blue pixels. However, the MR algorithm predicts the signal coverage at a low resolution, because only the average signal power of each cluster is predicted and the power will be assigned to each receiver within that cluster, while the NN algorithm predicts the signal coverage for all the receivers.

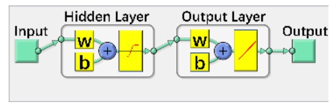

Fig. 2. NN structure.

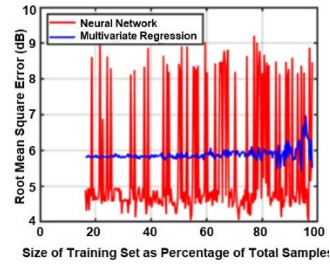

Fig. 3. RMSE comparison.

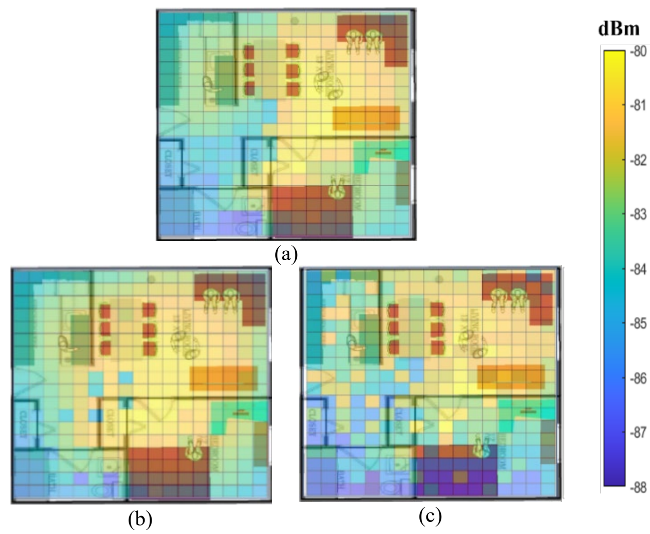

Fig. 4. Average signal power corresponding to the position of transmitter, obtained by: (a) Wireless InSite; (b) NN; (c) MR. The transmitter should be placed in the bright yellow regions for best overall coverage.

IV. Conclusions

This paper has presented a machine learning NN algorithm to efficiently predict 5G signal coverage to optimize transmitter deployment in an indoor apartment environment with high resolution, and compared it to the statistical MR method. The NN algorithm yields accurate predictions with the knowledge of only 20% of the data collected from the simulation software. This implies that the transmitter deployment strategy can be inferred quickly and accurately.

References

[1] H. Huang, et al, ''Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system,'' *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8549–8560, Sep. 2018.

[2] Z. Li, et al., "Impact of wall penetration loss on indoor wireless networks," *IEEE Wirel. Propag. Lett.*, vol. 20, pp. 1888–1892, 2021.

[3] T. Imai, et. al, "Outdoor-to-indoor path loss modeling for 0.8 to 37 GHz band," in *European Conf. Antennas Propag. (EuCAP)*, Davos, Switzerland, Apr. 2016, doi: 10.1109/EuCAP.2016.7481469.

[4] L. Azpilicueta, et al., "A ray launching-neural network approach for radio wave propagation analysis in complex indoor environments," *IEEE Trans. Antennas Propag.*, vol. 62, no. 5, pp. 2777–2786, 2014.