Machine Learning Based Indoor Path Loss Model

Debstuti Biswas(1), Sayan Mondal (2), Sujoy Mondal (3), and Abhirup Das Barman*(4)
(1) University of Arizona, Tucson, AZ 85721, United States
(2) Institute of Radiophysics and Electronics, University of Calcutta, West Bengal 700009, India
(3) RCC Institute of Information Technology, West Bengal 700015, India,
(4) Institute of Radiophysics and Electronics, University of Calcutta, West Bengal 700009, India

Abstract

The present work proposes wireless indoor path loss modelling and prediction by employing the synergy of real-time experiments and advanced machine learning techniques. Empirical path loss exponent factor for certain distance from the transmitter is obtained utilizing the minimum mean square error (MMSE) method from the experimental Received Signal Strength Indicator (RSSI) and path loss model. A hybrid Machine Learning (ML) model named as "AttentiveRSSI-LSTMNet" is implemented to predict correct Received Signal Strength (RSSI) in the indoor environment beyond the distance by the experimental setup. ESP32 micro-controller is designed for efficient transmitting and receiving data in our indoor wireless communication experiment. Distance versus outage probability is found out for varying transmitting powers. Our results have captured the proposed model's efficiency in assessing varied signal propagation characteristics in indoor settings and are able to provide precise predictions of RSSI beyond the measured distance using the proposed ML techniques.

1. Introduction

Superior methods of estimating path loss of the indoor channel are essential in enhancing characteristics of a network such as latency and data metrics. By identifying the general behavior of the channel, it will be possible to ensure a correct allocation of bandwidth and power and therefore a reduction of the operating costs. Crucial improvements have been made in machine learning based RSSI prediction. Linear regression and artificial neural networks (ANN) are utilized in N. Raj's work for RSSI prediction in indoors and taking into account the effects on precision [1]. A. A. S. AlQahtani et al makes use of machine learning for position prediction in Wi-Fi 2.4 GHz wireless networks wherein its application in indoor location-based services is discussed [2].

In this work, a novel hybrid model is proposed that not only analyzes signal variability to environmental changes but also enhances RSSI prediction accuracy beyond the measured distance.

2. Experimental Setup

Figure 1 depicts a wireless network setup comprises a transmitter node (Node-01), a receiver node or mobile station (M-1 device), and a local Wi-Fi router (WSN-AP) with Wi-Fi signal frequency 2.4 GHz. Transmitted data collected from the receiver is sent to a laptop. The distance between the transmitter and receiver is in meter range.

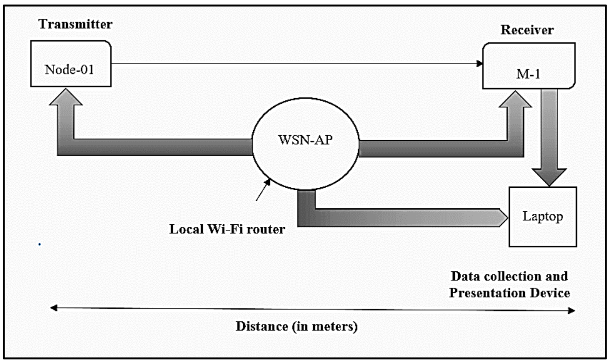

Figure 1. Wireless setup for measuring indoor path loss exponent

Figure 2 shows the complete floor mapping details of rooms, furniture, and walls of the corridor of the departmental building of Institute of Radio-physics and Electronics (IRPE) where we conduct indoor experiment. Signal propagation is influenced by these structures and on the measured RSSI values.

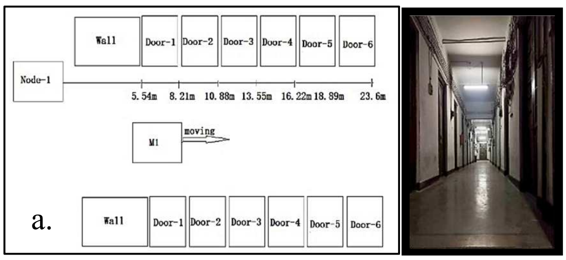

Figure 2. (a) Floor Mapping of Indoor Environment **(b)** Corridor of building (indoor environment)

ESP32 micro-controller shown in Fig.3 is designed for efficient transmitting and receiving data for our indoor wireless communication experiment.

Figure 3. ESP32 micro-controller based **(a)** Transmitting Node and **(b)** Mobile Station or Receiver Node (*components used are same for both*)

3. Methodology

Received signal power at a distance d can be written as

$$P_r(dBm) = P_t(dBm) + K(dB) - 10\gamma log_{10}\left(\frac{d}{d_0}\right) \quad (1)$$

Where, P_r, is the received power, P_t is the transmitted power, γ is the path loss exponent, d is the transmission distance, d_0 is a reference distance for the antenna's far-field region wherein antenna's radiation pattern is stable and radiating, and K is a unitless constant given by [3]:

$$K\,(dB) = 20\,\log_{10}(\lambda/\,4\pi d_0) + 10\,\log A + G_t + G_r \quad (2)$$

Where G_t and G_r are the gains of the transmitting and receiving antennas, respectively, λ is the wavelength of the signal and A is the effective area of ESP32 antenna.

3.1 MMSE Equations

We define $M = P_r/P_t$ (3)

$$M_{model}(d_i) = K(dB) - 10\gamma log_{10}\left(\frac{d}{d_0}\right) \quad (4)$$

By measuring the received power at select distances, we write mean square equation (MSE) to determine γ.

$$F(\gamma) = \sum_{i=1}^{N}\left[M_{measured}(d_i) - K(dB) + 10\gamma log_{10}\left(\frac{d}{d_0}\right)\right]^2 \quad (5)$$

In order to find the optimal γ, we differentiate the above equation with respect to γ i.e. $\frac{\partial F(\gamma)}{\partial_\gamma}$ and set it to zero. Solving this equation returns the value of γ [3].

Substituting all the known values used in the experiment yields **γ** ≈ 1.39. Using the γ value the model RSSI values are calculated. The model and the measured (Table-I) RSSI values are plotted with distance. The result is shown in Figure 1. Figure 1 suggests that the modelled RSSI values, using a path loss exponent of 1.39, provide a smooth and consistent estimate of signal strength variation over distance. On the other hand, the measured RSSI values exhibit substantial fluctuations due to the indoor environment that incorporates factors such as signals reflecting off walls, absorption coefficients of walls, and presence of various obstacles such as human movements and reflective surfaces [4].

TABLE-I: Measured vs Modelled Path Loss (dB) with Distance (m)

Index (i)	**Distance (d in m)**	**Measured RSSI (dB)**	**Modelled RSSI (dB) at γ=1.39**
0	3.12	-58	-62.51084343
1	5.79	-56	-66.26249511
2	8.46	-62	-68.5634676
3	11.13	-66	-70.22780602
4	13.8	-74	-71.53250872
5	16.47	-73	-72.60571383
6	19.51	-87	-73.63350017
7	21.85	-76	-74.32081087
8	24.52	-75	-75.02034447
9	27.19	-80	-75.64750235
10	29.86	-76	-76.2158666

In Figure 1, the solid line represents the theoretical values computed using the MMSE technique and extrapolated up to 60 meters. The dotted line with markers showcases the collected actual values presented up to 29.86 meters gathered throughout the study.

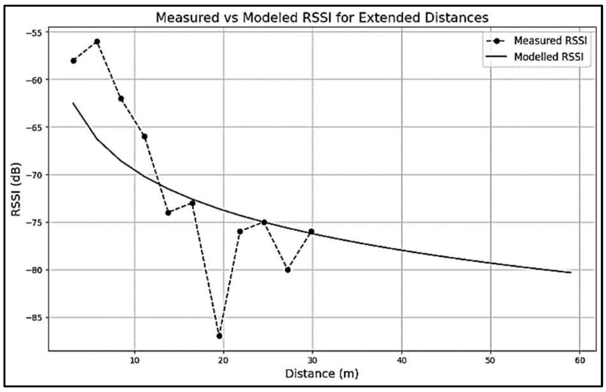

Figure 1. Measured RSSI up to 29.86m and Model RSSI Values for extended range of 60 meters.

3.2 Standard Deviation

The standard deviation of measured and modelled RSSI values can be given by [3]:

$$\sigma_{dB} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[M_{measured}(d_i) - M_{model}(d_i)]^2} \quad (5)$$

Thus, substituting γ=1.39 and all other known values, in (9) and using (11), standard deviation of the fluctuation is found to be $\sigma_{dB}(indoor)$ = **5.9849 dB.** Modelled RSSI values are calculated using (3) and (4). The calculated RSSI values are shown in Table I.

3.3 Outage Probability

The probability that the received power Z^2 falls below the predefined threshold (P_{min}) and it follows Gaussian statistics [3]. So we have,

$$P_{outage} = 1 - Q\left(\frac{P_{\min\_dBm} - \mu_{dB}}{\sigma_{dB}}\right) \quad (6)$$

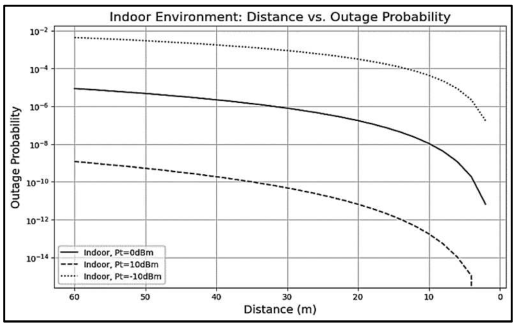

Figure 2. Outage probability in indoor environment for different transmit powers within 60-meters range.

Where, μ_{dB} is the average received power can be obtained from (4) and σ_{dB} is obtained from Eq. (5) and Q is similar to error function defined in [5]. We analysed the outage probability for different transmit powers (P_t = 0 dBm, P_t = 10 dBm, P_t = -10 dBm) at different distances as shown in Fig. 2. Outage probability worsens as we move away from the transmitting node.

4. Machine Learning for RSSI Prediction

The relationship between modelled RSSI with distance is not found linear as shown in Fig. 4. Non-linear models like Artificial Neural Networks (ANN) can be used for treating the non-linear relationship. A machine learning model is trained using the measured RSSI values which in turn can be used to predict RSSI values beyond 30m.

4.1 Model Architecture

We implement of our hybrid machine learning model named **"AttentiveRSSI-LSTMNet"** which is utilized to enhance RSSI prediction beyond the distance covered by the measured data. This model integrates both Dense Neural Networks (DNN) used for handling non-sequential and complex features within data and Recurrent Neural Networks (RNN) for capturing sequential dependencies over the data. The model architecture is given in Figure 3.

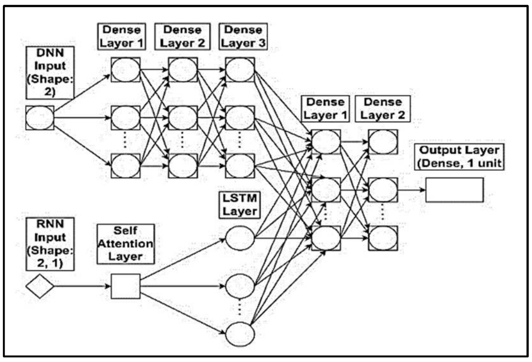

Figure 3. Diagram of the hybrid model "AttentiveRSSI-LSTMNet" architecture (with the *circle* labelled *"DNN Input"* representing the DNN component and the *rhombus* labelled as *"RNN Input"* representing the RNN component)

The first component which is the DNN Component has an input layer of shape (2), which means 2 features per sample as input into the network, followed by three Dense Layers with 32-256 units with ReLU activation. The RNN Component has an input layer with shape (2-1) meaning 2 time-step sequence with one feature per time step suitable for passing sequence data to LSTM layers. It has an initial self-attention layer with the attention width 15 and sigmoid activation, and LSTM layer with number of units ranging from 32 to 256 and default tanh and sigmoid activations respectively. The Concatenate Layer (consisting of 1 dense layer) is used for merging the outputs of the DNN and RNN components, succeeded by another Dense Layer with units ranging from 32 to 256 and ReLU activations respectively. The final output layer consists of a single unit for RSSI prediction.

4.2 Data Preparation

This model is trained with measured RSSI data acquired at fixed positions in an experimental environment, up to 29.86 meters. Hence, the dataset consists of a long set of empirical indoor RSSI measurements stored in .csv files at select distances ranging from 3.12m to 29.86m, consisting of both linear and non-linear data.

The data was split into: **Training Data:** 80% of the dataset, and **Testing Data:** 20% of the dataset. Using **'StandardScaler'** features were structured to ensure uniform contribution to model performance.

4.3 Training and Hyperparameter Optimization

Algorithm for training and hyperparameter optimization process is given below:

Step 1: Split the dataset, dividing it into 80% for training and, 20% for testing and validation. The consolidated dataset has 506 samples from 11 CSV files.
Step 2: Normalize features to have a zero mean and one standard deviation using Standard Scaler.
Step 3: Scale the feature dimensions if they are relevant to the RNN input.
Step 4: Stack three more dense with ReLU activation function.
Step 5: Integrate a self-attention layer and an LSTM layer within the RNN input stream.
Step 6: Connect the DNN and RNN parts through a concatenate operation.
Step 7: Train the model for 100 epochs and a batch size =5.
Step 8: Use the stochastic gradient descent algorithm to achieve the minimum MSE.
Step 9: Compare the parameters of the model in question with the test set, using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and R-squared (R^2) score.

4.4 Performance Metrics

The performance of our hybrid ML model was evaluated using several metrics between the predicted and real RSSI values in terms of MAE, MSE, RMSE and R^2 score [6]. Table II represents the significant performance improvements offered by our hybrid model over conventional models, demonstrating its superiority in terms of higher predictive accuracy and lower error rates.

TABLE-II: Comparative Performance Analysis of Traditional Models versus Our Hybrid Model “AttentiveRSSI-LSTMNet”

Model	MAE	MSE	RMSE	R^2
Basic LSTM [7]	1.10	1.25	1.12	0.92
DNN [8]	1.05	1.20	1.10	0.93
Hybrid DNN-LSTM [9]	0.95	1.05	1.02	0.94
Linear Regression [10]	1.25	1.56	1.25	0.85
Random Forest [11]	1.10	1.34	1.10	0.88
CNN-LSTM [12]	0.95	1.20	1.10	0.90
AttentiveRSSI-LSTMNet	**0.039**	**0.001**	**0.039**	**0.999**

The advantages of our hybrid model are as follows:

- Showcases superior performance metrics as evident in the R^2, MSE, MAE scores and Maximum Error score of **0.05**.
- Concatenates temporal sequence modelling and high-level feature extraction.
- Can handle different data scales and is adaptable to various environments.
- Self-attention mechanism focuses on significant parts of the data sequence, enriching the model’s vitality.

4.5 Comparison between Modelled and Predicted RSSI

Considering the previously presented experimental and theoretical results, the black line in Fig. 4 presents the modelled RSSI values as discussed in section 3. The proposed hybrid ML model’s RSSI values are described by dashed lines. The continuous line deviation is due to the standard deviation (5.98 dB) between simulated and modelled values as a result of reflections and obstacles. The ML model-predicted RSSI closely matches with the modelled one at the 30-60m plane where no experimental data were collected.

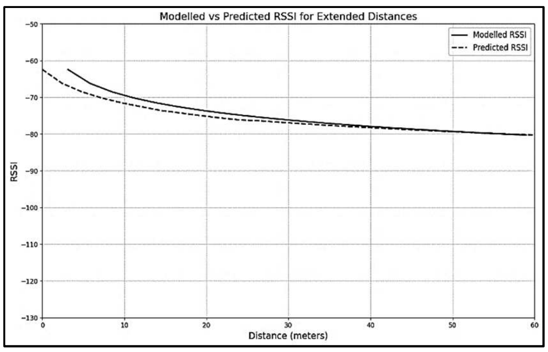

Figure 4. Modelled RSSI and Predicted RSSI (obtained from our hybrid model) up to an extended distance of 60 meters.

5. Conclusion and Future Directions

This work successfully models an indoor path loss using MMSE by using the measured RSSI values to find the outage probability. RSSI beyond experimental distances is predicted by a hybrid machine learning model which closely match with modelled RSSI values. Hence the ML approach can be used to find path loss exponent factor in the indoor environment which can find application in the field of radars, self-driving cars, the Internet of Things, and fifth-generation networks.

6. References

1. N. Raj, “Indoor RSSI Prediction using Machine Learning for Wireless Networks,” 2021 International Conference on COMmunication Systems & NETworkS (COMSNETS), Jan. 2021, doi: 10.1109/comsnets51098.2021.9352852.

2. A. A. S. AlQahtani and N. Choudhury, “Machine learning for location prediction using RSSI on Wi-Fi 2.4 GHz frequency band,” 2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Oct. 2021, doi: 10.1109/iemcon53756.2021.9623104.

3. A. Goldsmith, “Wireless communications,” 8 August 2005.

4. Richard, Ken, Martin, Richard, ‘’Building material and propagation’’ 14th sept.2014.

5. Haykin, S. (2008), “Communication systems”, John Wiley & Sons.

6. D. Chicco, M. J. Warrens, and G. Jurman, “The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation,” PeerJ. Computer Science, vol. 7, p. e623, Jul. 2021, doi: 10.7717/peerj-cs.623.

7. S. Hochreiter and J. Schmidhuber, “Long Short-Term memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

8. Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

9. M. Al-Ghamdi, A. A.-M. Al-Ghamdi, and M. Ragab, “A hybrid DNN multilayered LSTM model for energy consumption prediction,” Applied Sciences, vol. 13, no. 20, p. 11408, Oct. 2023, doi: 10.3390/app132011408.

10. G. James, D. Witten, T. Hastie, and R. Tibshirani, An Introduction to Statistical Learning: with Applications in R. 2021. [Online].

11. L. Breiman, “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, Jan. 2001, doi: 10.1023/a:1010933404324.

12. Shi, Xingjian, et al. "Convolutional LSTM network: A machine learning approach for precipitation nowcasting." Advances in neural information processing systems 28 (2015).