Comparison of Statistical and Deep Learning Path Loss Model for Motherboard Desktop Environment

Jinbang Fu, Prateek Juyal, Erik J. Jorgensen, and Alenka Zajić Senior Member, IEEE
Georgia Institute of Technology, Atlanta, GA 30332 USA

Abstract—In this paper, the path loss of the THz channel in motherboard desktop environment has been characterized and modeled by both statistical (mixture distributions) and deep learning (MLP) models. The performance of the two different model classes are compared and results show that mixture models captures the randomness of the channel by matching the PDF of measured path loss, which means that the statistical model can adapt to the changing environment. However, for the complex yet static motherboard desktop environment, the deep learning model outperforms the statistical models since it can also precisely describe the hidden patterns due to resonant modes and signal propagation in the static environment.

Index Terms—path loss model, THz communications, propagation, measurements.

I. INTRODUCTION

In the past few years, the field of communications has witnessed the development and deployment of 5G cellular network as well as the proposition and application of WiFi 6. These new techniques boost the data rate of wireless communications to several tens of gigabits per second. The promised high speed and low latency of the network fosters the growth of IoT applications. Since 2019, the number of active IoT devices increased from 7.7 to 10 billion and is expected to reach 25 billion by 2030 [1]. The explosive growth of IoT devices again requires better rack-to-rack, device-to-device, and chip-to-chip communications to meet the increasing demand for data transfer and storage. To further improve the data rate and reduce latency, attention has shifted to terahertz frequency for larger available bandwidth. Especially for chip-to-chip communication, more antennas can be integrated on a single chip to provide additional links since THz frequency communication allows smaller antennas and reduced spacing.

To construct a THz intra-chip communication system, on-board THz wireless channel measurements have been performed considering several possible scenarios [2]. Also, the characterization of desktop and nettop-sized metal cavities has been conducted and both traveling wave and resonant modes have been found inside the cavity [3], [4]. For the path loss at THz frequency, some models have been proposed that combine the propagation loss and the molecular attenuation [5]. Additionally, considering the rapid change and ultra-wide bandwidth of THz channel, a Gamma mixture model for path loss has been proposed in [6]. Based on the characterized propagation mechanism inside the cavity, the mean path loss has been modeled as the superposition of traveling loss, resonant-based power variation, and the loss due to the misalignment between transceivers for in-cavity scenarios [4], [7].

This paper continues our work in [7] and models the path loss in a motherboard desktop environment over the measured frequency band at different transceiver heights. Measurements have been performed over the motherboard inside a metal cavity. The in-cavity measured path loss has been characterized by both statistical and deep learning models in this work. We compare performance of the different models and show that both models have good approximation to the characterized THz channel. We find that the deep learning model outperforms the statistical models as it captures hidden patterns due to resonant modes and signal propagation in the static environment.

The rest of the paper is organized as follows. Section II presents the details of mixture models. Section III presents the design of our MLP (multilayer perceptron) network. Section IV shows the measurements, model verification and comparison. Finally, Section V makes concluding remarks.

II. PATH LOSS MODEL VIA MIXTURE DISTRIBUTIONS

Path loss always fluctuates over the frequency spectrum. This fluctuation is described as the shadowing across the frequency range and can be treated as a random process described by a Gaussian or Gamma distribution. However, unlike the wireless channel at low frequency, THz signals are very sensitive and their vulnerability to air molecules yields different performance across sections of the ultra-wide frequency band. Hence, it is not sufficient to model the extremely wide band with a single distribution [6]. To fit the actual THz wireless channel, a mixture distribution better approximates the characteristics of each sub-channel. Here, we model the path loss at different heights with both Gamma and Gaussian-mixture models and compare performance with an MLP neural network.

A mixture model consists of two types of parameters, the component weights and the component parameters (μₗ and σᵢ²) which are component mean and variance. A mixture model with L components can be expressed as [6]:

\[ f_c(x) = \sum_{l=1}^{L} \rho_l f_l(x; \mu_l, \sigma_l^2), l = 1, 2, \ldots, L, x > 0, \rho_l > 0, \]  \tag{1} 

where \( \rho_l \) represents the mixture proportion or weight of the \( l^{th} \) component which satisfies the conditions that \( 0 < \rho_l < 1, \forall l = 1, 2, \ldots, L, \) and \( \sum_{l=1}^{L} \rho_l = 1. \) For the Gaussian
mixture model, \( f_l(x|\mu_l,\sigma_l^2) \) follows the Gaussian distribution where \( f_l(x|\mu_l,\sigma_l^2) = \frac{1}{\sqrt{2\pi}\sigma_l^2} \exp\left(-\frac{(x-\mu_l)^2}{2\sigma_l^2}\right) \), whereas \( f_l(x|\mu_l,\sigma_l^2) = x^{\alpha_l-1} \exp(-x/\beta_l)/(\beta_l^\alpha_l\Gamma(\alpha_l)) \) with \( \mu_l = \alpha_l\beta_l \) and \( \sigma_l^2 = \alpha_l\beta_l^2 \) for the Gamma mixture model.

For a given \( L \), the EM (expectation maximization) algorithm is commonly used to determine maximum likelihood parameters of the mixture model. EM is an iterative algorithm which cycles between two steps. The first step is known as the expectation (E)-step which determines the membership coefficient with the expectation of component assignments for each data point. That is \( \forall n,l, \phi_{n,l} = \frac{\rho_l f_l(x_n|\mu_l,\sigma_l^2)}{\sum_{l=1}^L \rho_l f_l(x_n|\mu_l,\sigma_l^2)} \), where \( \phi_{n,l} \) is a membership coefficient that represents the probability that the \( n^{th} \) input value \( x_n \) is generated by the \( l^{th} \) component with the the distribution \( f_l(\cdot) \). The second step is the maximization (M)-step in which the calculated membership coefficients are used to update the parameters \( \rho_l, \mu_l, \) and \( \sigma_l^2 \) by maximizing the expectations calculated in E-step. That is \( l \in \mathbb{N} \),

\[
\dot{\mu}_l = \frac{\sum_{n=1}^N \phi_{n,l} x_n}{\sum_{n=1}^N \phi_{n,l}},
\]

\[
\dot{\mu}_l = \frac{\sum_{n=1}^N \phi_{n,l} x_n}{\sum_{n=1}^N \phi_{n,l}},
\]

\[
\hat{\sigma}_l^2 = \frac{\sum_{n=1}^N \phi_{n,l} (x_n - \dot{\mu}_l)^2}{\sum_{n=1}^N \phi_{n,l}}.
\]

where \( N \) is the number of input values and \( \rho_l, \dot{\mu}_l, \) and \( \hat{\sigma}_l^2 \) are the updated weight, mean, and variance of the \( l^{th} \) component, respectively.

III. PATH LOSS MODEL VIA MLP NETWORK

This section introduces a method to model the path loss in motherboard desktop environment with an MLP neural network. Statistical models like the mixture models described in Section II generate a good estimate of the mean and variation of measured path loss which approximate the channel with a corresponding probability density function (PDF). However, some patterns of the path loss generated from the physical arrangement of the environment cannot be described by statistical models nor explained from the propagation of EM waves when the environment is complicated. To handle this problem, we introduce an MLP neural network which can precisely capture those inconspicuous patterns.

MLP is a feed-forward artificial neural network which consists of multiple layers of nodes: an input layer at front, an output layer at the end, and at least one hidden layer in between. Each node in the hidden layers acts as an artificial neuron and is generated with a nonlinear activation function. Except for the input nodes, the value of each node in the MLP is given by an affine function with the nodes in previous layer as the inputs. A nonlinear activation function is applied after the affine function for each hidden unit to prevent the network from collapsing into a linear model. The neurons in an MLP can be trained with the back-propagation technique to allow optimization through gradient descent. An MLP can be used to solve regression problems and can distinguish data that is not linearly separable due to its many layers and nonlinear activations. In our measurement setup, the path loss was measured in a motherboard desktop environment over a 300 - 312 GHz frequency band with different transceiver heights over the motherboard. Consequently, we train the model to predict the path loss over the frequency band of interest for any given transceiver heights. To achieve this goal, a seven layer decoder-like MLP network was constructed as shown in Fig. 1. Transceiver height is the only feature input to the MLP network and the output layer generates 801 output activations corresponding to the measured frequency band that was sampled with 801 points. Five hidden layers are inserted between the input and output layers. These five hidden layers consist of 20, 50, 100, 200, and 400 nodes, respectively. The LeakyReLU activation function is applied at the output of each hidden-layer, where LeakyReLU is a ReLU- (rectified linear unit) based activation function that outputs itself for positive values and has a small slope (0.01 for our case) for negative values. When training an MLP network, the nodes in hidden layers may take values with strong variation along the layers of the network, across different nodes in the same layer, and over time due to the updating model parameters. This strong variation could prevent convergence of the network and may require active adjustment of the optimizer's learning rate for the compensation. To handle this problem, we applied batch normalization after the affine transformation but before the activation functions. As suggested by its name, batch normalization normalizes the inputs by subtracting the mean and dividing by the standard deviation estimated from the current mini-batch for each training iteration. Then a coefficient and an offset are applied to the normalized inputs. When training our network, we use the “Adam” variant of the stochastic gradient descent optimization algorithm which estimates the first- and second-order moments of the gradient via an exponential moving average to update parameters.

![Fig. 1. Architecture of our seven layer decoder-like MLP network.](image-url)
Fig. 2 compares the measured path loss at selected heights $h = 0.3, 0.6, 1.5, 2.1, 3.3, 4.2,$ and $4.5$ cm with the Friis prediction. Due to the reflection on the motherboard and distribution of resonant modes inside the cavity, as shown in Fig. 3, there is at least a 7 dB (65 - 72 dB) variation of the average measured path loss with respect to the transceiver height, $h$. A strong constructive effect can be observed when $h = 0.6$ and $4.2$ cm, the measured path loss varies around 65 dB. Also, a strong destructive effect can be seen when $h = 3.9$ cm. For most other transceiver heights, the measured path loss varies between 67 and 69 dB around the Friis prediction. Also, the measured path loss oscillates differently over the frequency band at different heights. The oscillation in measured path loss is due to the resonant modes and the interference between the line-of-sight and multipath components generated from the scattering on the motherboard.

IV. ANALYSIS AND MODEL VERIFICATION

In this section, we modeled the THz wireless channel in motherboard desktop environment with the statistical and deep learning methods described in Section II and III. Then we compared and analyzed the performance of these models relative to a baseline. To model the THz channel, we first performed measurements of the path loss at across a range of transceiver heights. An aluminum metal cavity of size $27.5 \times 27.5 \times 10$ cm was fabricated as shown in Fig. 2a. The top and bottom walls of the cavity were made with two square aluminum plates with side length of $27.5$ cm separated by $10$ cm from each other by two stacked $5$ cm nylon standoffs at each corner, and the other four sides of the cavity wrapped with aluminum foil. As shown in Fig. 2b, a motherboard was put inside the cavity and supported on the bottom wall by three stacked brass $14$ mm (with extra $6$mm threaded) standoffs at each corner of the motherboard. These brass standoffs were affixed to the motherboard with four screws and to the bottom plate with four additional standoffs. As pictured in Fig. 2a, the fabricated cavity was put between the transceivers with antennas aligned horizontally during the measurements. Antennas of the transceivers were inserted into the cavity through two square openings cut on the opposite sidewalls of the cavity. The two $42.08$ mm-long horn antennas were fixed to a $19.1$ cm distance between Tx and Rx. Measurements were performed over the motherboard by simultaneously varying the height of transceivers, $h$, from 0 to $4.5$ cm with the step size of $0.3$ cm. The parameter $h$ refers to the distance between the bottom edge of the horn and the motherboard, and there is $4.46$ mm between the phase center and bottom edge of the horn.

The oscillation across the frequency is a shadowing of path loss and can be described with a random process. To characterize the shadowing of path loss at different heights, the Gamma and Gaussian mixture models as described in section II are applied. Figure 4 compares the PDF of measured path loss with Gamma and Gaussian-mixture distributions described by equation (1) at $h = 0.3, 2.1, 3.9,$ and $4.5$ cm. For the Gamma-mixture model, the corresponding number of mixed Gamma distributions, $k_n$, are $9, 16, 4,$ and $13$, respectively. The number of mixed Gaussian distributions, $k_g$, are set to be $9, 9, 9,$ and $15$, respectively for the Gaussian-mixture model. Both $k_g$ and $k_n$ are chosen to be the smallest possible while still achieving an R-squared goodness of fit to be greater than $97%$. We note that fitting to the measurements can reach a higher accuracy with larger $k_g$ or $k_n$. It can be observed from the plots that both Gamma and Gaussian-mixture distributions have good fit to the PDF of the in-cavity measured path loss. The curves drawn by the Gamma-mixture model are smoother and closer to the PDF of measured path loss, while the Gaussian-mixture model provides more details.

The THz wireless channel has also been modeled with an MLP network and fed with measured and synthetic data for training. The measured path loss can be derived by subtracting...
the S21 from the sum of antenna gains and the thru loss. Before the measurements, we first characterized the thru loss on the transceiver devices ten times. Then, as mentioned previously, we measured the antenna gain at sixteen heights \( h \) from 0 to 4.5 cm with the increment of 0.3 cm. During these measurements, we collected ten S21s at each of the \( h = 0, 0.6, 1.2, 1.8, \) and 2.4 cm heights in addition to three S21s from each of the remaining heights. We combined 10 thru loss and 83 S21 measurements to make 830 samples in total (100 each at \( h = 0, 0.6, 1.2, 1.8, \) and 2.4 cm, and 30 each at the other eleven heights). From these 830 measurements, we randomly picked two from the data collected at each height and applied the data augmentation described in Section III on the 32 total selected measurements. After the augmentation, we discarded the data whose value was set to None, resulting in 5510 synthetic path loss in total for training. From the rest of the path loss measurements, three were randomly picked from the measurements at each height and interpolated with Akima interpolation. After discarding the None values, we got 1350 augmented data for the validation. The remaining 750 path loss measurements were used for testing.

Before training, we first normalized the training set (for both features and labels) by subtracting its mean and dividing by its standard deviation. This normalization plays well with the optimizer since it scales the parameters to a similar level. The validation and testing sets were normalized with the mean and standard deviation of the training set. The MLP network was trained for 1000 epochs. During each epoch, the network iterates over random batches of 256 samples and updates the values of its parameters with each batch. Figure 5 compares the behaviors of training, validation, and testing data sets’ MSE (mean square error) loss during the training process. It can be seen from the plot that the MSE loss of training and validation data sets are always lower than that of the testing data set. The MSE loss dropped quickly in the first 50 epochs of training since the model parameters are initialized randomly. The MSE losses of training, validation, and testing data sets converged to 0.08, 0.08, and 0.11, respectively, after 500 epochs.

![Fig. 4. Gamma (k_\gamma mixtures) and Gaussian (k_\nu mixtures) mixture models for the PDFs of in-cavity measured path loss with:](image)

(a) \( h = 0.3 \) cm: \( R^2(k_\gamma = 9) \): 97.28% \( R^2(k_\nu = 9) \): 97.13% \( R^2(k_\nu = 9) \): 97.06%

(b) \( h = 2.1 \) cm: \( R^2(k_\gamma = 16) \): 97.62%, \( R^2(k_\nu = 9) \): 97.06%

(c) \( h = 3.9 \) cm: \( R^2(k_\gamma = 4) \): 97.02%, \( R^2(k_\nu = 9) \): 97.10%

(d) \( h = 4.5 \) cm: \( R^2(k_\gamma = 13) \): 97.01%, \( R^2(k_\nu = 15) \): 97.04%.

![Fig. 5. MSE loss of training, validation, and testing sets over 1000 epochs.](image)

Figure 6 compares the MLP network, Gamma-mixture model, and Gaussian mixture model with randomly selected path loss measurements at \( h = 0.3, 2.1, 3.9, \) and 4.5 cm. As shown in Fig. 6a and Fig. 6b, the predictions of Gamma and Gaussian-mixture models generated from the modeled PDFs match well with the mean values and variances of measured path loss, but not with the hidden oscillatory patterns. Examples of these hidden patterns are circled in red for each model, and are due to the resonant modes and signal propagation in the complex static environment. The MLP network outperforms the statistical models by matching these hidden patterns, as shown in Fig. 6c. The MLP not only captures the noise variation of the channel, but also matches well with the hidden patterns, as circled in red. This is because those hidden patterns are characteristics of the channel which are identical in different measurements at the same height. As opposed to the mixture models, the MLP network can learn those patterns from the training data.

A baseline was derived to visualize the performance of mixture models and the MLP network by averaging the measured path loss of the training set at each height. The MSE losses between the baseline and the mixture models, designed MLP network, and randomly picked measured path loss at \( h = 0.3, 2.1, 2.4, 3.3, 3.9, \) and 4.5 cm were calculated and shown in Fig. 7. As shown in the plot, comparing with the baseline, the MSE losses of mixture models are varying between 0.3 and 1.5, and the Gamma-mixture model performs slightly better than the Gaussian-mixture model. The MLP network outperforms the mixture models since its MSE losses at different heights are lower than 0.2. Additionally, the variation of its MSE losses over \( h \) is minimal, which demonstrates the robustness of the MLP model. The MSE
losses of most randomly picked path loss are less than 0.05 while two are relatively large (0.22 and 0.38 at $h = 2.4$ and 4.5 cm, respectively). This means that the variation of different measurements at each height is small, proving that most patterns of the measured path loss are characteristics of the channel and can be learned by the MLP model. For that reason, the deep learning model outperforms the statistical generative models in the complex static environment.

V. CONCLUSIONS

This paper presented path loss modeling for the THz channel in motherboard desktop environment. Measurements were performed by varying the heights of transceivers over the motherboard inside an artificial desktop cavity. The measured path loss has been modeled by mixture models and an MLP network. The results show that mixture models match well with the PDF of the measured path loss, but the MLP network outperforms the mixture models since it can also capture the hidden frequency-variant patterns due to the resonant modes and the signal propagation in the static environment.

REFERENCES