A Data-Driven Approach to Localization for High Frequency Wireless Mobile Networks

Marcus Z. Comiter, Michael B. Crouse, and H. T. Kung
John A. Paulson School of Engineering and Applied Sciences
Harvard University
Cambridge, Massachusetts 02138
Email: marcuscomiter@g.harvard.edu, mcrouse@seas.harvard.edu, kung@harvard.edu

***Abstract*—The use of high frequency millimeter wave (mmWave) bands at 28 GHz or higher will be a defining characteristic of next-generation wireless networks, such as 5G and 802.11ad networks. However, communicating over these high frequencies bands often requires directional antennas on base stations to dynamically align their beams with mobile nodes. To quickly align antennas, we propose a data-driven deep neural network (DNN) approach to localize mobile nodes using lower frequency spectrum. Our methods require fewer than 30 real-world sample locations to learn a model that can localize a mobile node to the required 5G indoor sub-meter accuracy. We demonstrate with real-world data in indoor and outdoor experiments that this performance is achievable, even in multipath-rich environments. Further, we show via simulation that the proposed DNN approach is robust to noise and collinearity between antenna arrays. Our primary contributions are: (1) a novel structure for a deep neural network that reflects the impact base station location has on node localization, (2) a quantized loss function for neural network training that improves localization accuracy and reduces the amount of training data needed, and (3) a procedure for using synthetic data to reduce the required number of real-world measurements needed for training the data-driven localization model. Our real-world experiments show that the use of synthetic data can improve localization accuracy by over 3x.**

Keywords*—millimeter waves, 5G, localization, mobile networks, machine learning, synthetic data for training.

I. Introduction

Future wireless networks, including 5G cellular and 802.11ad networks, will rely on high frequency millimeter waves (mmWaves) to deliver substantially increased data rates and bandwidth. A consequence of using high frequency signals is the need for directional antennas to combat their higher attenuation, requiring that communicating endpoints align their beams [1]. In order to provide accurate and fast beam alignments, mobile nodes can be localized such that base stations can correctly select their beam sectors. The emerging 5G standard calls for sub-meter localization, highlighting its importance for enabling efficient and accurate beam alignment [2]. This paper proposes highly accurate and low-latency data-driven localization methods for mmWave beam alignment specifically suited for mobile networks.

Localization is a important task that has been well-studied under a wide variety of assumptions, technologies and techniques. However, the use of localization information within a mobile network protocol introduces hard constraints on latency, accuracy, and available sensing technology. 802.11 base stations already contain lower frequency (e.g., 2.4 or 5 GHz) antennas for use in lower data rate applications but can also be used as sensors for localization. Fortunately, lower frequency signals are better suited for localization as their omni-directional beams allow for more efficient localization as higher frequency, directional beams would require an expensive antenna sector scanning process. However, a major challenge with utilizing lower frequencies is their complex multipath behavior, especially in indoor environments due to the large variety of materials and objects that may be present.

We propose a fully data-driven deep neural network (DNN) approach to localization to address the challenges associated with using low frequency signals. This approach learns the received signal patterns associated with a target environment to overcome the difficulties introduced by complex multipath effects, while delivering highly accurate localization with a small amount of latency. To reduce the amount of real-world measurement data required for training the model, we show that synthetically created data can be effectively used to augment real-world measurements. Further, we show that a relatively small neural network model (e.g., with four hidden layers, each with 50 to 200 neurons) is sufficient and that calculating the location can be performed in under a millisecond via a feed-forward operation.

Our methods differ from a traditional DNN, as we tailor the structure (topology) of our model to address the specific challenges associated with the underlying task, such as collinearity between base stations, fading, and multipath. This domain-specific neural network approach can increase localization accuracy while simultaneously minimizing the amount of real-world samples that must be collected for training the model. By altering black-box machine learning methods for the domain, we can localize nodes at higher accuracy while simultaneously limiting the overly onerous amounts of data collection that would render the method impractical for real-world use. Our main contributions are as follows:

1) We introduce a **structured neural network model** and a **novel quantized loss function** to address collinearity between base stations, complicated multipath effects, and fading.
2) We introduce a procedure for using **synthetic data** created from limited real-world data, reducing the need for time-consuming data collection, and therefore making our proposed methods practical for real-world use cases.
3) We introduce a **simulator** for empirical evaluation of localization accuracy and the effectiveness of using

978-1-5090-3019-2/17/$31.00 ©2017 IEEE

synthetic data in model training.

4) We collect indoor and outdoor **real-world data** using an array of software defined radios, and demonstrate on this data that our methods can achieve **sub-meter localization fidelity**.

II. Related Work

The authors of [3] present methods for tracking users through walls using frequency modulated carrier waves (FMCW), or a chirp signal. This method relies on the users to be moving in order to separate their reflections from the static background, and requires a windowed set of measurements for accurate background subtraction.

The 802.11ad Amendment III proposes to perform beamforming and alignment through a procedure called Sector Level Sweep (SLS), an in-band search that scans available antenna sectors, and then refines via the Beam Refinement Process (BRP). The authors of [4] introduce Blind Beam Steering (BBS) to reduce the amount of scanning by utilizing a lower frequency, 2.4 GHz, and computing an angle-of-arrival using algorithms such as Multiple Signal Classification (MUSIC) to reduce the amount of scanning. While our method also uses lower frequencies, our data-driven methods can learn multipath effects from observed data. Further, by calculating a position rather than just the angle-of-arrival, our method can be directly useful for antenna selection as well as mmWave MAC algorithms for users at various locations.

ArrayTrack uses a set of base stations with antenna arrays for localization using angle-of-arrival measurements [5]. However, it requires as many as eight antennas on each base station in order to reduce the error when determining the angle-of-arrival and improving robustness to the multipath effects in indoor wireless environments. To achieve the similar accuracy, our base stations utilize only four antennas or less per array, a result of the adopted data-driven approach.

Data-driven, deep neural networks for performing localization has been previously studied under multiple indoor scenarios [6], [7], [8]. However, their accuracy does not meet the required sub-meter fidelity specified for the 5G standard for indoor environments (see [9] for a study on misalignment impact on channel quality). Our method improves upon the localization accuracy by modifying the neural network structure (topology) to suit the underlying localization task, as well as introducing a method to synthetically create training data. Our use of synthetic training data draws inspiration from image recognition with deep convolutional neural networks where training data is often augmented with synthetically transformed versions of the original data [10]. Synthetic data generation for localization of wireless nodes is crucial for reducing the burden of collecting real-world training data and improving overall accuracy.

III. Goals and Challenges for Localization

To motivate the challenges involved in using millimeter wave (mmWave) for wireless networks, we illustrate a working scenario in Figure 1. Several dual mode base stations are distributed throughout the environment, and are assumed to be capable of operating at two different frequencies: low frequencies for localization (such as 2.4 GHz and 916 MHz, as in our real-world experiments) and higher frequencies for high data-rate mmWave communications (such as 28 GHz, 38 GHz, or 60 GHz). The localization goal is to meet the sub-meter accuracy called for in the emerging 5G wireless standard for indoor deployments [2]. We now discuss the challenges for accurate localization introduced by collinearity between base stations, multipath, and meeting the low-latency requirements.

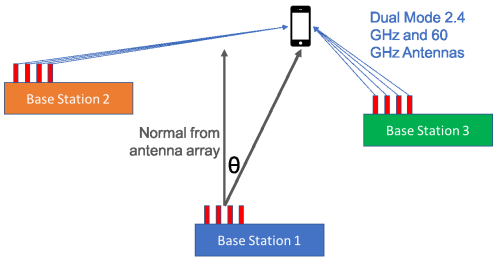

Fig. 1. Model of localization setup for mmWave networking, where each base station is equipped with an antenna array. The diagram illustrates the manner in which angle-of-arrival is measured, denoted by θ.

A. Collinearity

One challenge for localizing a mobile node is the existence of collinear regions between base stations. We define a *collinear region* between two base stations as the region along a straight line connecting the two base stations. The collinear region is especially relevant in the context of localization, as the location of any node located along this line cannot be uniquely determined from the angles-of-arrival between the node and the two base stations or the phase offsets at the two base stations. Collinear regions introduce a *one-to-many relationship*, as one input (e.g., angles-of-arrival or phase offsets) maps to multiple locations within the collinear region, making accurate localization difficult.

Figure 4 demonstrates the consequences of collinearity in practice for a simplistic DNN model that we discuss in Section IV-A. The model predicts all nodes within the collinear region as being at the midpoint of the collinear region. This stems from the fact that minimizing the loss for a one-to-many mapping is accomplished by predicting the output as the average of the outputs of all data points whose inputs are the same.While having data from more than two base stations would seem to resolve these regions, we find that conventional machine learning models have trouble resolving these regions, or require a much larger amount of training data.

B. Multipath

An additional challenge for accurate localization is the propagation characteristics of low frequency signals. Specifically, low frequency waves exhibit complex multipath behaviors as they reflect and diffract off of walls and other objects differently based on their physical characteristics (shape and material). Complex multipath behaviors can introduce large variations when estimating information such as angle-of-arrival, as they alter the phase or time of flight measured at each of the base stations. Therefore, multipath must be carefully accounted for or the predicted location can be very inaccurate. Several advancements have been proposed to combat multipath complexity; however, they are often limited by the number of

physical antennas or the number of potential paths within the environment [5], [11].

In our data-driven approach, as long as regions within the environment with large multipath effects are sampled, these key characteristics can be captured, or learned, by the model. We will show that even in the face of large multipath effects, only a few samples are required to sufficiently capture these effects, as discussed in Section VI-A, and that additional training data can be generated synthetically to further mitigate multipath effects, as discussed in Section IV-C.

C. Latency and Interference

In order to use localization for beam alignment purposes, the latency for obtaining a mobile node's location must be extremely low. This is complicated by the fact that many next generation applications, such as Virtual Reality and Augmented Reality, will have rapidly moving nodes operating within close proximity of one another. Given this, localization methods for beam sector selection must be sufficiently fast and not interfere with other nodes. Large latency in beam alignment and can delay system initialization, while intermittent latency (e.g., MAC ARQ) resulting in packet loss due to misaligned beams can disrupt higher layer protocols, such as by causing TCP timeouts, reducing of throughput [12]. Finally, we note that blind in-band beam scanning (e.g., the SLS approach in the 802.11ad standard) in high frequencies incurs delays.

IV. Localization Methods

In this section, we introduce the models and their domain-specific modifications for localizing mobile nodes. We first describe our Structured Multi-Layer Perceptron (SMLP), a modified deep neural network (DNN) that improves the localization accuracy and robustness for nodes within collinear regions. We next present a novel quantized loss function, the Epsilon Invariant Loss Function (EILF), which provides improved accuracy and allows smaller amounts of training data to be used to reduce model overfitting. Finally, we describe a principled method for generating a large amount of synthetic training data from a small set of real-world data. This increases localization accuracy by reducing the potential for data-driven models to overfit.

A. The Structured Multilayer Perceptron (SMLP)

We first present the Structured Multilayer Perceptron model (SMLP) as a domain-specific method for addressing the localization of mobile nodes within regions of collinearity. These modifications allow the model to mitigate the collinear problems that occur in regions between base station pairs. Without these modifications, it is challenging for conventional neural networks, such as Multi-layer Perceptrons (MLPs), to accurately localize mobile nodes in collinear regions with a small number of training samples, e.g. tens of samples.

We now detail the construction of the SMLP, which is illustrated in Figure 2. The SMLP is constructed of two types of layers: a Lower Pair Network (LPN) and an Upper Connection Network (UCN). The LPN consists of a combination of sub-models, each of which are fully connected MLPs. Each sub-model is trained independently of one another and only on a single pair of base stations. Once trained, the output layer

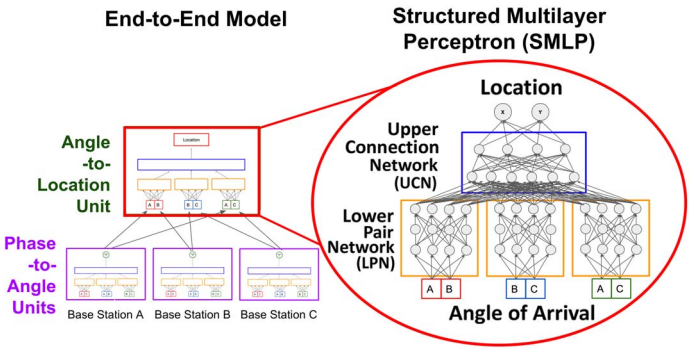

Fig. 2. A diagram of our proposed Structured Multilayer Perceptron (SMLP) model (right), which separates base stations pairs in the Lower Pair Network (LPN) and then combines them in the Upper Connection Network (UCN), and is a building block for the End-to-End Model (left). The SMLP infers the location of a mobile node in terms of its (x, y)-coordinates.

(location prediction) is removed and the last layer of the sub-models in the LPN are concatenated, used as input for the model in the UCN. The UCN is then trained to produce the final predicted location for each data point. For the wireless localization scenario, the inputs are angles-of-arrival, and the output is mobile node location. By combining the output of the multiple complementary LPN models at the UCN, the SMLP is able to resolve the collinear regions that the lower layer sub-models are unable to resolve by themselves. Intuitively, the lower layers of the model provide a denoising function by operating in a manner similar to an ensemble machine learning method. The output of each sub-model for points within the collinear regions will always be a denoised, constant value that is easy for the upper layer to combine with the input from the other sub-models to predict an accurate output location.

Additionally, we also use raw phase offsets captured from an antenna array as input, instead of angles-of-arrival, to calculate the location of a mobile device. To do this, we utilize the same SMLPs, but stack them on top of one another to create an End-to-End model, as shown on left of Figure 2. The bottom layer consists of a set of SMLPs that take phase offsets as input and produce angles-of-arrival as output. This output is then used as input to an upper layer SMLP, which takes the computed angles-of-arrival as input, and produces a final localization as output. In the real-world results presented in Section VI, we show that this End-to-End model produces accurate localization using raw phase offset information alone.

B. Epsilon Invariant Loss Function

We next present the Epsilon Invariant Loss Function (EILF), a novel quantized loss function that improves localization accuracy and allows for less data to be used in training. Figure 3 illustrates that when using localization for determining beam alignment, the level of accuracy only needs to be within the range of the antenna beamwidth, i.e. wider beams require less accurate localization for correct alignment. The EILF loss function is defined in Equation 1, where the level of accuracy is specified by ϵ.

$$L(y, \hat{y}) = \begin{Bmatrix} 0 & \text{if } ||y - \hat{y}||_2^2 \leq \epsilon \\ ||y - \hat{y}||_2^2 & \text{if } ||y - \hat{y}||_2^2 > \epsilon \end{Bmatrix}. \quad (1)$$

Equation 1 specifies that when the error between the

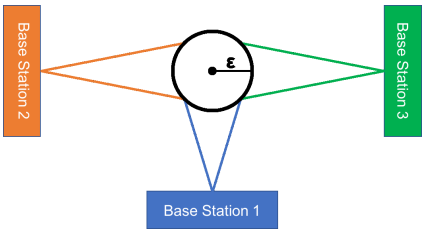

Fig. 3. The Epsilon Invariant Loss Function captures the fact that while the beamwidth is narrow, it still covers epsilon region rather than just a point.

predicted location $\hat{y}$ and the true location y is greater than ϵ, the loss $||y - \hat{y}||_2^2$ is kept as a standard $\ell 2$ loss function. Note, the user can configure ϵ according to the physical properties of mmWave antenna arrays, a wider beamwidth means ϵ should be larger. Otherwise, the loss is set to zero as the predicted error is within the required accuracy for the given antenna beamwidth. By reducing the loss to 0 in these cases, further rounds of training are instead focused on areas with larger error. The EILF causes the learner to focus on points with significant error by not penalizing the model's performance in areas in which localization is sufficiently accurate. We show in Section V-B that EILF allows significantly fewer points to be used in training without a loss in localization accuracy.

C. Synthetic Data

For real-world practicality, it is critical to limit the amount of user-sourced measurement necessary for accurate localization; therefore, it is advantageous to be able to utilize synthetically generated data to augment a small set of real-world samples. Our detailed experiments show that are slight variations in measurements between days, which result in poor localization accuracy on models trained between days. We observe that these variations from the mean measurement of each location are approximately Gaussian distributed. To formalize this, we use an Anderson-Darling statistical test to see if the differences from the mean in measurements over days follows a Gaussian distribution. On the "Lounge" dataset we collected, we find that a test statistic of 0.18, below the 0.74 critical value at the 5% significance level, which supports our hypothesis regarding the distribution.

Based on this observation, we introduce a method to generate synthetic data from a limited amount of collected real-world data, shown in Algorithm 1. Given a set of real-world measurements, a user specified amount of synthetic data to generate $\mathbf{n}$, where $\mathbf{n}$ is a multiplicative factor of how many times the amount of collected data is to be synthesized, and an amount of variance $\mathbf{v}$ used in generating the data, the synthetic data is generated as follows: for each data point $\mathbf{d}_i$ in the real-world collected data, $\mathbf{n}$ points are sampled from a Gaussian distribution with mean $\mathbf{d}_i$ and variance $\mathbf{v}$. All of the generated points, along with the original real-world data, are then combined, called the "Real + Synthetic Data Set" $\mathbf{D}'$.

Intuitively, the synthetic data allows the model to deal with noise and other sources of variations in measurement data that may not be fully captured in the limited training data. Without the synthetic data, the model is susceptible to overfitting the peculiarities of the limited amount of training data. As shown empirically in Section VI-A, synthetic data use can reduce error by over a factor of three.

Algorithm 1: Synthetic Data Generation Algorithm

Input: Training Data Set $\mathbf{D}$, amount of synthetic data generation $\mathbf{n}$, amount of variance $\mathbf{v}$
Output: Augmented Training Set $\mathbf{D}'$
$\mathbf{D}' \leftarrow D$
for $\mathbf{d}_i \in \mathbf{D}$ **do**
  for $\mathbf{j} \in 1 \ldots \mathbf{n}$ **do**
    $\mathbf{d}'_j \leftarrow$ Sample from Gaussian$(\mathbf{d}_i, \mathbf{v})$
    $\mathbf{D}' \leftarrow \mathbf{D}' \cup \mathbf{d}'_j$

V. Simulation Results

We now discuss the simulated results to validate our proposed methodologies. We first describe the simulator we have built for these purposes, which we have released publicly as a community resource. We next present results using the simulator under a number of different conditions. For all results in this section, we use a virtual setup with three base stations, and evaluate on a held-out test set of 2500 points canvasing the entire virtual space. All results on localization accuracy are reported in terms of Median Squared Error (Median SE). In the rest of this section we assume that the virtual space corresponds to a 6m x 6m physical region. Note that Median SE under 1 meets the 1m fidelity goal.

A. Simulator

To evaluate our proposed Structured Multi-layer Perceptron and Epsilon Invariant Loss Function, we created a simulator for generating large quantities of training data under different scenarios. The simulator generates a virtual space with mobile nodes and base stations. It then generates different types of raw inputs for the machine learning models (e.g., angles-of-arrivals) based on the location of the base stations and mobile nodes. The simulator generates either time of flight or angle-of-arrival to each of the base stations given a virtual mobile node's position and can be configured via initialization files. The current simulator can be configured with the following options: distribution of mobile locations (uniform grid, random), number of mobile locations, dimensionality of simulation (2D, 3D), and number and location of base stations. We also incorporate several noise and filtering modules, including Gaussian noise, Angle Dependent Noise, base station outages, and uniform noise. The code for the simulator, as well as the real-world data sets collected, can be found online [13].

B. Collinearity

We now evaluate the performance our SMLP for addressing the collinearity problem. We train a SMLP model with 2 layers in the Lower Pair Network (LPN) with 500 and 50 neurons per layer, respectively, and 2 layers in the Upper Connection Network with 200 and 50 neurons per layer, respectively. Note that a commodity accelerator capable of executing 4500 GFLOPs can execute the $(3 \times 3 \times 500 \times 50) + (1 \times 50 \times 200 \times 50) = 725,000$ multiply and addition operations in

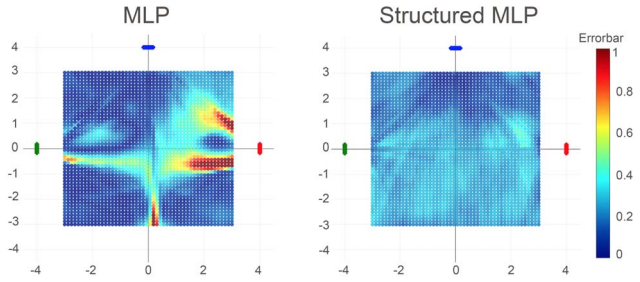

Fig. 4. When simulating data without noise, the non-structured MLP is unable to resolve the collinear regions while our proposed Structured MLP model can. There are three base stations surrounding each region, donted by small colored bars.

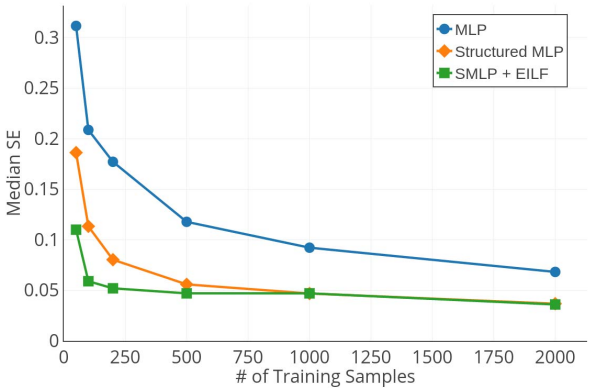

Fig. 5. Using simulation data without noise, we show the improvement in localization accuracy of the Structured MLP (SMLP) with and without EILF over a non-structured MLP model. The SMLP with EILF can localize nodes with the same accuracy using only 5% of the training data needed for the non-structured MLP and 20% for the SMLP.

0.02 milliseconds. As a point of comparison, we train a (not-structured) fully connected MLP model with 4 hidden layers and 500, 50, 200, 50 neurons per layer, respectively. Figure 4 shows the error in localizing points within the simulated space with three base stations for the MLP and SMLP models. The MLP model has large error in the collinear regions between base stations (denoted by the red regions), but the SMLP model is able to resolve these regions. We additionally train an MLP with more neurons per layer such that the total number of neurons is comparable to that of the SMLP, and find that this does not have a marked impact on improving its performance.

We next examine the Median SE achieved by each of the models in localizing a grid of 2500 data points within the simulated space. Figure 5 shows the Median SE achieved by the SMLP and non-structured MLP models for different amounts of training data, ranging from 50 to 2000 points. We find that for any amount of training data used, the SMLP outperforms the MLP in terms of Median SE. Intuitively, this is due to the fact that given two points in different positions but within the same collinear region, the UCN in the SMLP is able to use the output from the two LPN sub-models to resolve the collinear region problem, while the MLP, even with the same number of layers, is not able to do so.

Finally, we examine the use of the Epsilon Invariant Loss Function (EILF), as introduced in Section IV-B. In this set of experiments, we use the SMLP model, but utilize loss function with EILF. We set $\epsilon = 0.01$ in the loss function, which corresponds to a reasonable amount of allowed error when considering the beamwidth of mmWaves [14]. Figure 5 compares the accuracy of the SMLP model with and without EILF (as well as to the non-structured MLP model). We see that with small amounts of training data EILF reduces localization error significantly even beyond the gains made by the SMLP, achieving over two times improvement over the non-structured MLP. With a large amount of training data, the performance of the two SMLP models begin to converge because it allows the SMLP model without EILF to reduce error below the chosen ϵ value. without EILF to reduce error below the chosen ϵ value. EILF uses $\epsilon = 0.01$ for all of our remaining results as it provides an improvement in accuracy in all cases.

C. Noise

We now present results under two types of noise, Additive Gaussian White Noise (AGWN) and Angle Dependent Noise. AGWN can be caused by slight antenna movement, inaccuracies in measurement data, equipment mis-calibrations, and similar phenomena. To generate the AGWN data, n points are sampled from a Gaussian distribution with mean μ and standard deviation σ, where n is the number of data points in the dataset, and then add them to the generated input data. In these experiments, μ is set to 0 and σ is varied.

Figure 6 shows the impact of the AGWN model on both the proposed SMLP method and the non-structured MLP for different amounts of noise and different amounts of training data. The noise is sampled from Gaussian distributions with $\mu = 0$ and $\sigma = \{0.01, 0.02, 0.05\}$, translating to average amounts of 2.9 degrees, 5.7 degrees, and 14.4 degrees of noise in terms of angles of arrival, respectively. We find that the SMLP outperforms the non-structured MLP at all settings of noise and amounts of training data, with the gap between the models increasing as σ decreases. Further, with small to medium levels of noise ($\sigma \leq 0.02$) with as few as 100 training points, the 1 meter fidelity goal is easily reached. Finally, we find that at even high levels of noise ($\sigma = 0.05$), given a sufficient amount of training data (250 points), the SMLP model is able to meet the 1 meter fidelity goal, and with smaller amounts of training data, performance suffers only marginally (error rising from 0.99 median meters at 250 training points to 1.03 median meters with only 100 training points).

The second type of noise we consider is Angle Dependent Noise. When using phase offset information, measurements become noisier as the angle between the base station array and mobile node becomes large, i.e. far from the normal, denoted by θ in Figure 1. More specifically, when the mobile node is aligned in front of the base station antenna (i.e., close to the normal from the antenna array), there is little noise compared with when the mobile node is oriented, for example, 90 degrees from the normal of the antenna array. We define a nonlinear function in Equation 2, where the angle between the mobile node and base station is $\mathbf{x}$, $\mathbf{k}$ controls for the magnitude of the nonlinearity, and $\mathbf{j}$ controls for the base amount of noise (i.e., noise when the angle is 0 degrees). Figure 7 illustrates Angle Dependent Noise present in real measurements we collected in our outdoor field experiments and our nonlinear noise model

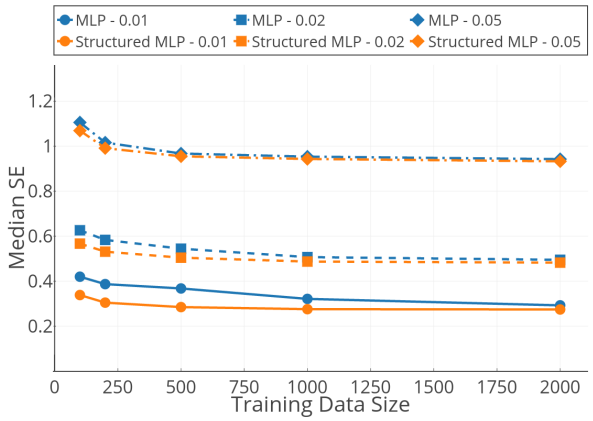

Fig. 6. Under Additive Gaussian White Noise (AGWN) under three standard deviations 0.01, 0.02, and 0.05; the Structured MLP (SMLP) outperforms the non-structured MLP model.

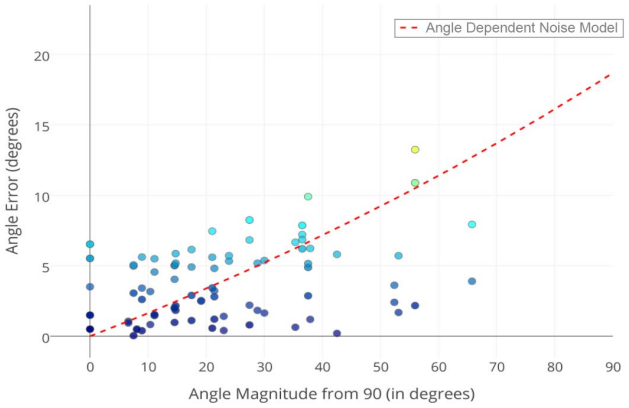

Fig. 7. Our Angle Dependent Noise model (red line) reflects real-world data captured from our outdoor field experiments (points).

described in Equation 2 plotted in red ($k = 1$ and $j = 4$). To generate data with Angle Dependent Noise, we first generate noise from a Gaussian distribution and then multiply by the nonlinear factor according to the location of the point, as shown in Equation 3.

$$a(x) = e^{|k(x-90)|} - j \tag{2}$$

$$a(x) \times \mathcal{N}(\mu = 0, \sigma) \tag{3}$$

To evaluate our models under Angle Dependent Noise, we consider varying levels of Gaussian noise, $\sigma = \{0.01, 0.02, 0.05\}$, multiplied by the nonlinearity defined in Equation 2 with $k = 1$ and $j = 4$. Figure 8 shows the impact of Angle Dependent Noise with increasing Gaussian noise and varying amounts of training data on localization accuracy. Overall, we find that the amount of localization error under the Angle Dependent Noise model is less than that under the AGWN model as the Angle Dependent Noise model primarily affects those points at a greater angular offset while the AGWN model affects all points equally.

The SMLP significantly outperforms the MLP with small and medium amounts of noise at all amounts of training data, with the gap between the models increasing as σ decreases. The SMLP model at $\sigma = 0.02$ achieves about the same error as the non-structured MLP model at $\sigma = 0.01$ with small amounts of training data. Additionally, with as few as 100 points, the 1

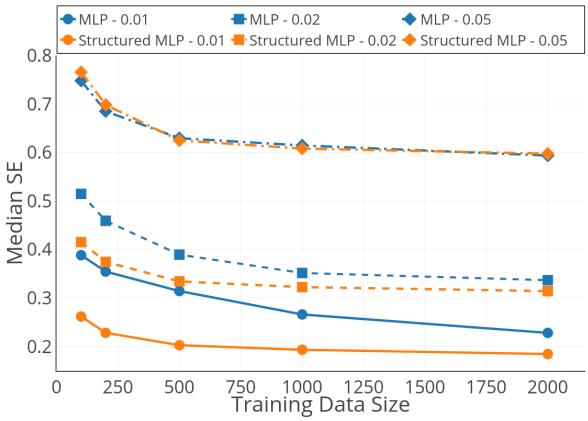

Fig. 8. Using simulation data with Angle Dependent Noise and AGWN, we show the improvement in localization accuracy of the Structured MLP (SMLP) as compared to the non-structured MLP model.

Dataset	Real (Phase)	Real + Synthetic (Phase)
Outdoor	0.74	0.33
Lounge	1.17	0.37
Classroom	1.20	0.76

TABLE I. MEDIAN SE LOCALIZATION ACCURACY ON REAL-WORLD EXPERIMENTS IN THREE ENVIRONMENTS ACROSS MULTIPLE DAYS.

meter fidelity goal is easily reached (with a localization error of 0.51m, 0.65m, and 0.88m, for $\sigma = 0.01$, 0.02, and 0.05).

VI. REAL-WORLD EXPERIMENTS

We now present real-world experiments performed in multiple environments, each with varying levels of multipath. Our experimental setup contains a collection of receivers, a transmitter, and a reference node. We use N200 User Software Radio Peripheral (USRP) software defined radios operating at 916 MHz. The receiver USRPs are controlled by a portable Dell Inspiron 410 connected via a Netgear gigabit switch. The transmitter (mobile node) sends a string of a thousand 1s that is captured by all of the receiver USRPs (base stations). The reference point is placed directly in front of the base station so the phase offsets can be used for calibration [15]. We use three separate test environments: one limiting multipath to at most a single bounce, and two with uncontrolled amounts of multipath. We use a 2D setup as we lack enough equipment for dual polarity. We use a 5 by 6 grid of 30 location points measuring 416.5 cm x 266.4 cm, with each sub-grid measuring 83.3 cm x 66.6 cm. The three base stations are placed 90 cm from the midpoint of three sides of the rectangular grid.

The first environment is an outside, open, grassy area with a simple bounce coming only from the grass beneath the USRPs. The second environment is an indoor classroom with tables moved out of the center in order to allow for exact placement of the transmitter and repeatable experiments. The third environment is an indoor lounge. We permit regular foot traffic to walk through the experimental area during the lounge experiment, resulting in a more realistic environment.

A. Results

We now present results on all three real-world environments. For each experiment, we collect 90 samples (3 samples

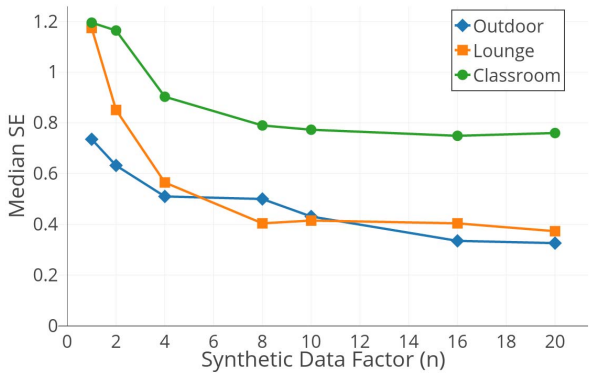

Fig. 9. Real-world localization accuracy increases with increasing amounts of generated synthetic training data for all our experiments, realizing greater than 2x improvement on average.

at each of the 30 points on the grid) per day for two days. The model is trained on data from one day, and tested on data from the other day. The second day measurements are completely held out from training and synthetic data generation. We use the End-to-End SMLP model described in Section IV-A, utilizing only the raw phase information as input to the model.

Results are shown in Table I comparing real and synthetic data training. The "Real Only" results use a model trained only on actual samples collected from our real-world experiments. The "Real + Synthetic" results use a model trained on samples collected on one day, as well as synthetic data generated from these samples using the procedure described in Section IV-C. We use a synthetic data generation factor of $\mathbf{n} = 20$, a variance of $\mathbf{v} = 0.03$, and generate synthetic data according to Algorithm 1. The addition of synthetic data reduces error by a factor of 2.3 on average over all real-world experiments. In the "Lounge" experiment, the addition of synthetic data reduces error by a factor of 3.2. In the "Outdoor", "Lounge", and "Classroom" experiments, the addition of synthetic data allows for localization accuracy of 0.57, 0.60, and 0.87 meters, respectively, well within the 1 meter fidelity goal.

Finally, we present results demonstrating how the amount of synthetic data used impacts performance. Figure 9 shows how model performance is affected by the amount of synthetic data used, with accuracy improving as more synthetic data is introduced. The lounge environment, which is the most complex, illustrates the largest improvement, reducing the error from well over a meter to less than half a meter. The returns from additional synthetic data begin to reduce as the noise begins to outpace the signal in the data.

VII. Conclusion

We have described an improved data-driven localization approach for narrow beam alignment for mmWave networks. Our real-world experiments and simulation results achieve the sub-meter localization accuracy called for by the 5G mmWave standard. Our proposed data-driven methods automatically learn an environment with only a small number of real-world measurements, (e.g. tens of measurements), even under multipath. We use three methods to achieve this result: 1) novel neural network structures specifically designed for the localization task, 2) a quantized loss function for improving accuracy and reducing the amount of training data needed, and 3) a procedure for generating synthetic data to augment real measurements needed for training a model. We show that our methods are robust under various noise models and address challenges of collinearity between base stations. Our improved localization performance will enable low-latency, accurate beam alignment for next generation wireless networks, and can be used in the future to inform higher layers in the network stack to improve access control and link utilization.

VIII. Acknowledgements

This work is supported in part by gifts from the Intel Corporation and in part by the Naval Supply Systems Command award under the Naval Postgraduate School Agreements No. N00244-15-0050 and No. N00244-16-1-0018.

References

[1] T. Nitsche, C. Cordeiro, A. B. Flores, E. W. Knightly, E. Perahia, and J. C. Widmer, "Ieee 802.11 ad: directional 60 ghz communication for multi-gigabit-per-second wi-fi [invited paper]," *IEEE Communications Magazine*, vol. 52, no. 12, pp. 132–141, 2014.

[2] N. Alliance, "5g white paper," *Next Generation Mobile Networks, White paper*, 2015.

[3] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller, "3d tracking via body radio reflections." in *NSDI*, vol. 14, 2014, pp. 317–329.

[4] T. Nitsche, A. B. Flores, E. W. Knightly, and J. Widmer, "Steering with eyes closed: mm-wave beam steering without in-band measurement," in *2015 IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2015, pp. 2416–2424.

[5] J. Xiong and K. Jamieson, "Arraytrack: A fine-grained indoor location system." in *NSDI*, 2013, pp. 71–84.

[6] G. Félix, M. Siller, and E. N. Álvarez, "A fingerprinting indoor localization algorithm based deep learning," in *Ubiquitous and Future Networks (ICUFN), 2016 Eighth International Conference on*. IEEE, 2016, pp. 1006–1011.

[7] J. Luo and H. Gao, "Deep belief networks for fingerprinting indoor localization using ultrawideband technology," *International Journal of Distributed Sensor Networks*, 2016.

[8] X. Wang, L. Gao, S. Mao, and S. Pandey, "Csi-based fingerprinting for indoor localization: A deep learning approach," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 763–776, 2017.

[9] M. Comiter, M. Crouse, H. T. Kung, J.-H. Tarng, Z.-M. Tsai, W.-T. Wu, T.-S. Lee, M. C. F. Chang, and Y.-C. Kuan, "Millimeter-wave field experiments with many antenna configurations for indoor multipath environments," in *2017 IEEE GLOBECOM The Fourth International Workshop on 5G/5G+ Communications in Higher Frequency Bands (5GCHFB)*. IEEE, 2017.

[10] M. Jaderberg, K. Simonyan, A. Vedaldi, and A. Zisserman, "Synthetic data and artificial neural networks for natural scene text recognition," *arXiv preprint arXiv:1406.2227*, 2014.

[11] R. Roy and T. Kailath, "Esprit-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989.

[12] S. J. Tarsa, M. Comiter, M. B. Crouse, B. McDanel, and H. Kung, "Taming wireless fluctuations by predictive queuing using a sparse-coding link-state model," in *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. ACM, 2015, pp. 287–296.

[13] M. Comiter and M. Crouse, https://github.com/KevinHCChen/wireless-aoa/, 2017.

[14] I. . working group *et al.*, "Ieee 802.11 ad, amendment 3: Enhancements for very high throughput in the 60 ghz band," 2012.

[15] H.-C. Chen, T.-H. Lin, H. Kung, C.-K. Lin, and Y. Gwon, "Determining rf angle of arrival using cots antenna arrays: a field evaluation," in *MILITARY COMMUNICATIONS CONFERENCE, 2012-MILCOM 2012*. IEEE, 2012, pp. 1–6.