MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

mmWave Wi-Fi Trajectory Estimation with Continuous-Time Neural Dynamic Learning

Vaca-Rubio, Cristian; Wang, Pu; Koike-Akino, Toshiaki; Wang, Ye; Boufounos, Petros T.; Popovski, Petar

TR2023-033 May 06, 2023

Abstract

We leverage standards-compliant beam training measurements from commercial-of-the-shelf (COTS) 802.11ad/ay devices for localization of a moving object. Two technical challenges need to be addressed: (1) the beam training measurements are intermittent due to beam scanning overhead control and contention-based channel- time allocation, and (2) how to exploit underlying object dynamics to assist the localization. To this end, we formulate the trajectory estimation as a sequence regression problem. We propose a dual- decoder neural dynamic learning framework to simultaneously reconstruct Wi-Fi beam training measurements at irregular time instances and learn the unknown dynamics over the latent space in a continuous-time fashion by enforcing strong supervision at both the coordinate and measurement levels. The proposed method was evaluated on an in-house mmWave Wi-Fi dataset and compared with a range of baseline methods, including traditional machine learning methods and recurrent neural networks.

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2023

© 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

MMWAVE WI-FI TRAJECTORY ESTIMATION WITH CONTINUOUS-TIME NEURAL DYNAMIC LEARNING

Cristian J. Vaca-Rubio[1,2], *Pu Wang*[1], *Toshiaki Koike-Akino*[1], *Ye Wang*[1], *Petros Boufounos*[1], *Petar Popovski*[2]

[1]Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA 02139, USA
[2]Department of Electronic Systems, Aalborg University, Aalborg, Denmark

ABSTRACT

We leverage standards-compliant beam training measurements from commercial-of-the-shelf (COTS) 802.11ad/ay devices for localization of a moving object. Two technical challenges need to be addressed: (1) the beam training measurements are intermittent due to beam scanning overhead control and contention-based channel-time allocation, and (2) how to exploit underlying object dynamics to assist the localization. To this end, we formulate the trajectory estimation as a sequence regression problem. We propose a dual-decoder neural dynamic learning framework to simultaneously reconstruct Wi-Fi beam training measurements at irregular time instances and learn the unknown dynamics over the latent space in a continuous-time fashion by enforcing strong supervision at both the coordinate and measurement levels. The proposed method was evaluated on an in-house mmWave Wi-Fi dataset and compared with a range of baseline methods, including traditional machine learning methods and recurrent neural networks.

Index Terms— WLAN sensing, Wi-Fi, 802.11ad/ay, localization, fingerprinting, beam training, dynamic learning.

1. INTRODUCTION

Wi-Fi fingerprinting is one of the popular approaches for indoor localization, driven by open firmware releases, WLAN domain knowledge for waveform calibration and preprocessing, as well as recent advances in deep learning-based feature extraction [1–3].

With commercial-of-the-shelf (COTS) Wi-Fi devices, one can fingerprint the following Wi-Fi measurements:

- coarse-grained received signal strength indicator (RSSI) [4];
- mid-grained beam training measurements at 60 GHz [5–11]
- fine-grained channel state information (CSI) at sub-7 GHz [12–15];

Refer to [10, Section II] for detailed discussions on all three types of Wi-Fi channel measurements. Traditional machine learning and advanced deep learning methods have been applied to all Wi-Fi fingerprinted measurements [16–21]. For instance, DeepFi exploits 90 CSI amplitudes from all the subcarriers at three antennas for feature extraction using an autoencoder architecture [12, 22]. More recently, a pretrained fusion network between the CSI at sub-7 GHz and the beam training measurements at 60 GHz was proposed for both localization and device-free sensing tasks [10]. Nevertheless, the majority of these approaches are *frame-based*; that is, the coordinate is inferred from the current Wi-Fi frame, without integration of past measurements or previous trajectory history.

The work of C.J. Vaca-Rubio was done during his visit and internship at MERL. He was also supported from European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant agreement No. 813999.

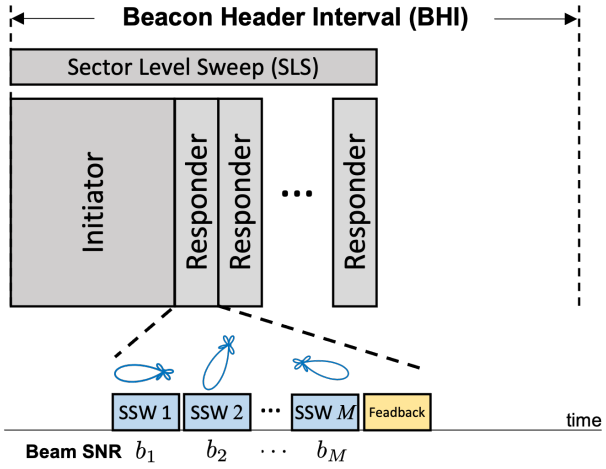


Fig. 1. Beam training measurements (Beam SNRs) during the mandatory Sector Level Sweep (SLS) in 802.11ad/ay standards.

On the other hand, *sequence-based* approaches take consecutive Wi-Fi frames as the input, and state estimation (e.g, Kalman filter-like approaches [23, 24]) and recurrent neural networks (e.g., GRU and LSTM [25]) can be applied for trajectory estimation with the RSSI [21] and CSI [26] at sub-7 GHz. However, the sequence-based formulation has NOT been applied to mmWave Wi-Fi localization due to the intermittent nature of the mid-grained beam measurement:

1) **Low beam training rate**: In Fig. 1, during the beacon header interval (BHI), mmWave Wi-Fi of 802.11ad/ay uses directional beacons for sector level sweep (SLS) to train both initiator/responder beampatterns for subsequent data transmission. This mandatory beam training results in significant overhead to the Wi-Fi network and it is desired to limit the number of directional beampatterns within a beacon and the total number of beacons, resulting in sparsely sampled beam measurements than Wi-Fi at sub-7 GHz.

2) **Irregular sample intervals**: Consider a scenario where the access point (AP) is the initiator and the users are responders. When the responder trains its (transmitting or receiving) beampatterns, a sequence of M sector sweep (SSW) frames is sent via different beampatterns to the initiator and the initiator can compute M corresponding SNRs, $b_1, b_2, \cdots, b_M$, within a responder channel time. When multiple users exist, each user needs to contend the next responder channel time and one contending user is randomly selected. As a result, the contention-based channel access results in irregularly beam SNR measurements at AP for a given user.

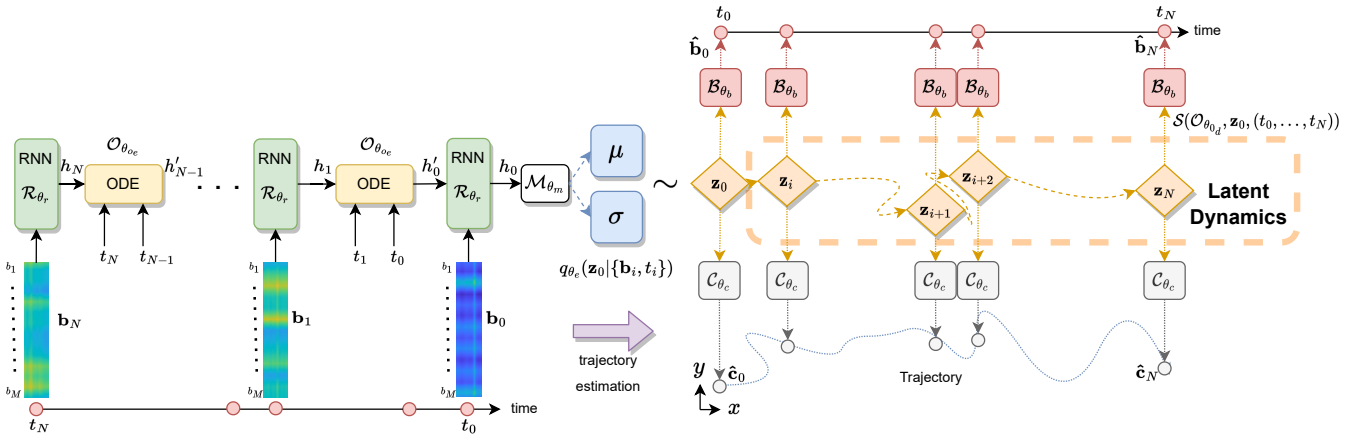


Fig. 2. Dual-Decoder Neural Dynamic Learning framework.

To address the above challenges and inspired by recent advances in neural ordinary differential equation (NODE) [27], this paper proposes a dual-decoder neural dynamic learning framework that learns a consistent ODE, via strong supervisions from both coordinate and measurements levels at two separate decoders, to describe unknown, continuous-time latent dynamics of the intermittently-sampled mmWave Wi-Fi measurements. Compared with the original single-decoder NODE, we show that the additional supervision at the coordinate level leads to strong performance gain. The proposed method was evaluated on an in-house mmWave Wi-Fi dataset and compared with a range of baseline methods.

2. PROBLEM FORMULATION

We formulate the indoor localization as a sequence regression of beam training measurements within a period of ΔT_w seconds for trajectory estimation. Specifically, stacking the M beam SNRs during one responder channel time t_i as $\mathbf{b}_i = [b_1, b_2, ..., b_M]^T \in \mathbb{R}^{M\times 1}$, the problem of interest is to utilize beam SNR measurements $\{\mathbf{b}_i\}_{i=0}^N$ at time steps $\{t_i\}_{i=0}^N$ with irregular sample intervals to localize the object,

$$\{\mathbf{b}_i, t_i\}_{i=0}^N \rightarrow \{\mathbf{c}_i\}_{i=0}^N, \quad s.t. \quad \Delta t_i = t_i - t_{i-1} \neq \Delta t_{i+1} \tag{1}$$

where $\mathbf{c}_i = [x_i, y_i]^T$ consists of corresponding two-dimensional coordinates (x_i, y_i) at t_i. This is illustrated in Fig. 2 where the trajectory estimation is to convert the set of beam SNRs $\{\mathbf{b}_i\}_{i=0}^N$ at intermittently-sampled steps $\{t_i\}_{i=0}^N$ (shown in the left bottom part) to the set of $\{\mathbf{c}_i\}_{i=0}^N$ over a continuous trajectory (shown in the right bottom part).

3. DUAL-DECODER NEURAL DYNAMIC LEARNING

We frame our method as a latent-variable model, that we denominate as dual-decoder neural dynamic (DDND). We present our framework in Fig. 2. Corresponding Fig. 2 from left to right, in the subsequent sections, we introduce the encoder structure for successive $\{\mathbf{b}_i\}_{i=0}^N$ sequences, the process of learning the latent trajectory, and the learning method empowered by strong supervision to enhance the continuous trajectory learning.

Notation: θ denotes the learnable parameters in neural networks. For simplicity, we use θ_e to denote the joint parameters of all the networks comprising the encoder. We use θ_{o_e} and θ_{o_d} to denote the parameters of the networks comprising the encoder and decoder ODE parts, respectively. We also use θ_r, θ_m, θ_b and θ_c to denote the parameters of the Recurrent Neural Network (RNN), the MLP that outputs the mean and standard deviation of the encoded signal, and the two linear decoders. $\mathcal{S}$ denotes an arbitrary ODE solver.

3.1. Waveform Temporal Information Encoding

Denote a sequence of beamSNR measurements within ΔT_w as $\{\mathbf{b}_i\}_{i=0}^N \in \mathbb{R}^{N\times B}$ and its corresponding coordinates $\{\mathbf{c}_i\}_{i=0}^N \in \mathbb{R}^{N\times 2}$. We present the input as a temporal sequence to encode the underlying dynamics of the variation of the mmWave Wi-Fi signal with regard to the physical trajectory. We obtain the encoded temporal information of every measurement by forwarding the described temporal inputs through an ODE-RNN network [27]. The ODE and RNN blocks are modeled as neural networks $\mathcal{O}_{\theta_{o_e}}(.)$ and $\mathcal{R}_{\theta_e}(.)$, respectively. When forwarding a beamSNR temporal sequence, we reverse the time sequence from t_N to t_0. In this way, the encoder network learns the approximate posterior at time t_0. These Neural ODE [28] blocks are used in the encoder network to model the evolution of the hidden states $\mathbf{h} \in \mathbb{R}^E$, where E denotes the dimension of the hidden states. This behavior is modeled in a continuous fashion $\mathbf{h}(t)$, as a solution to an ODE initial-value problem:

$$\frac{d\mathbf{h}(t)}{dt} = \mathcal{O}_{\theta_{o_e}}(\mathbf{h}(t), t), \tag{2}$$

$\mathcal{O}_{\theta_{o_e}}(.)$ defines the time-reversed evolution of the observed beamSNR states as the solution of an ODE:

$$\mathbf{h}'_{i-1} = \mathcal{S}(\mathcal{O}_{\theta_{o_e}}, \mathbf{h}_i, (t_i, t_{i-1})), \tag{3}$$

then, the hidden state is updated for each observation as a standard RNN update:

$$\mathbf{h}_{i-1} = \mathcal{R}_{\theta_r}(\mathbf{h}'_{i-1}, \mathbf{b}_{i-1}). \tag{4}$$

In our approach, we want to characterize $\mathbf{z}_0$ that represents the latent initial state of the encoded trajectory. For that purpose, the mean and the standard deviation of the approximate time-reversed posterior $q_{\theta_{o_e}}(\mathbf{z}_0|\{\mathbf{b}_i, t_i\}_{i=N}^0)$ are a function of the final hidden state of the encoder:

$$q_{\theta_e}(\mathbf{z}_0|\{\mathbf{b}_i, t_i\}_{i=N}^0) = \mathcal{N}(\mu_{\mathbf{z}_0}, \sigma_{\mathbf{z}_0}), \tag{5}$$

where

$$\mu_{\mathbf{z}_0}, \sigma_{\mathbf{z}_0} = \mathcal{M}_{\theta_m}(\mathcal{O}_{\theta_{o_e}}(\{\mathbf{b}_i, t_i\}_{i=N}^0)), \tag{6}$$

where $\mathcal{M}_{\theta_m}(.)$ is a neural network translating the last hidden state of the encoder into the mean and variance of the latent initial state $\mathbf{z}_0$.

3.2. Latent Dynamics

Once estimating the approximate posterior $q_{\theta_e}(\mathbf{z}_0|\{\mathbf{b}_i, t_i\}_{i=N}^0)$, the beamSNR variable-length input sequence $\{\mathbf{b}_i\}_{i=0}^N$ is encoded into a fixed-dimensional latent space embedding $\mathbf{z} \in \mathbb{R}^L$, where L denotes the dimension of the latent space. The latent trajectory is obtained by first sampling $\mathbf{z}_0 \sim q_{\theta_e}(\mathbf{z}_0|\{\mathbf{b}_i, t_i\}_{i=N}^0)$ from the estimated posterior. Then, on the decoder side, another ODE $\mathcal{O}_{\theta_{o_d}}$ is modeled as a neural network. During the training, $\mathcal{O}_{\theta_{o_d}}$ will learn the latent trajectory dynamics that relate the variation of the signal and the physical trajectory while during the forward pass, it will query the latent trajectory at the specified time instants. For these matters, $\mathbf{z}_0$ is used as the initial value for the ODE solver on the decoder side:

$$\begin{aligned} \mathbf{z}_0, ..., \mathbf{z}_N &= \mathbf{z}_0 + \int_{t_0}^{t_N} \mathcal{O}_{\theta_{o_d}}(\mathbf{z}_t, t)dt \\ &= \mathcal{S}(\mathcal{O}_{\theta_{o_d}}, \mathbf{z}_0, (t_0, ..., t_N)). \end{aligned} \tag{7}$$

Up to this part, the beamSNR input sequence has been decoded into the latent trajectory $\{\mathbf{b}_i\}_{i=0}^N \rightarrow \{\mathbf{z}_i\}_{i=0}^N$.

3.3. Dual Decoder

In order to guarantee a suitable latent trajectory learning dynamics, we propose to condition the learning by including two linear decoders in the decoder side: waveform reconstruction $\mathcal{B}_{\theta_b}(.)$, and trajectory regression $\mathcal{C}_{\theta_c}(.)$. These two heads will take as input the latent trajectory, in order to perform the reconstruction of the input signal

$$\hat{\mathbf{b}}_i = \mathcal{B}_{\theta_b}(\mathbf{z}_i) = \mathbf{W}_b \mathbf{z}_i + \mathbf{v}_b, \tag{8}$$

and its corresponding trajectory regression

$$\hat{\mathbf{c}}_i = \mathcal{C}_{\theta_c}(\mathbf{z}_i) = \mathbf{W}_c \mathbf{z}_i + \mathbf{v}_c, \tag{9}$$

where $\mathbf{W}_b$, $\mathbf{W}_c$ denotes the weight matrices and $\mathbf{v}_b$, $\mathbf{v}_c$ the bias vectors, respectively. Note that the input for both decoder heads is the predicted latent trajectory by the decoder ODE. Also, the proposed decoders use shared weights for the input sequences. In this way, we are imposing strong supervision for every time instant in the latent trajectory by using the real trajectory and the variation of the signal as conditions to modify the learning dynamics of the latent trajectory. This leads to an enhancement in learning the continuous dynamics of the trajectory from the latent space.

3.4. Dual-Decoder Neural Dynamic Loss

We propose to train in an end-to-end encoder-decoder structure to minimize the dual-decoder loss which is given by

$$\mathcal{L} = \left[\alpha \|\hat{\mathbf{c}}_0 - \mathbf{c}_0\|_1 + \frac{1}{N}\sum_{i=1}^{N} \|\hat{\mathbf{c}}_i - \mathbf{c}_i\|_1\right] + \beta \left[\frac{1}{N}\sum_{i=0}^{N} \|\hat{\mathbf{b}}_i - \mathbf{b}_i\|_1\right] \tag{10}$$

where the first term corresponds to the trajectory regression, and the second term to the waveform reconstruction. We also have two hyperparameters α and β to balance their importance during the learning. In the first term, we are weighting the first coordinate of the trajectory on its own with the α factor. This is done to enhance the trajectory learning, as we are solving an ODE initial-value problem, being the first point of the trajectory determinant for the trajectory regression.

3.5. Complexity Analysis

The time complexity of an ODE-RNN depends on the number of hidden units in the recurrent layer H and the number of time steps in the input sequence T. Then, the time complexity of the forward pass of the ODE-RNN can be approximated as $\mathcal{O}(TH^2)$. Similarly, for the linear layers, an input dimension N and an output dimension M will be expressed as $\mathcal{O}(NM + M)$. In this way, our proposed framework has approximately a time complexity:

$$\mathcal{O}\left(T(H_e^2 + H_d^2 + (L+1)(B+C))\right), \tag{11}$$

where the sub-indexes represent the encoder and decoder ODEs, respectively, L is the dimension of the latent space, and B and C are the dimensions of the output linear decoders, respectively.

Remark: Recurrent Neural Networks (RNNs) are not the best solution to learning irregularly-sampled dynamics. Although some tricks have been performed to address this problem, such as inputting the time difference information as a feature or computing an exponential decay between observations [29], they are still not modeling the underline continuous dynamic of an irregular time series data. This highlights the difficulty of intermittent-sampled data. That is why we propose this end-to-end structure that first tries to encode the continuous dynamics of the signal along the trajectory into a latent distribution. This latent distribution represents the encoded dynamics of the motion up to the first location of the trajectory, regardless of the intermittency due to the NODE. Modeling ODEs as neural networks in both the encoder and the decoder side enables learning the continuous behavior of the data and provides great flexibility to design the decoder side.

4. PERFORMANCE EVALUATION

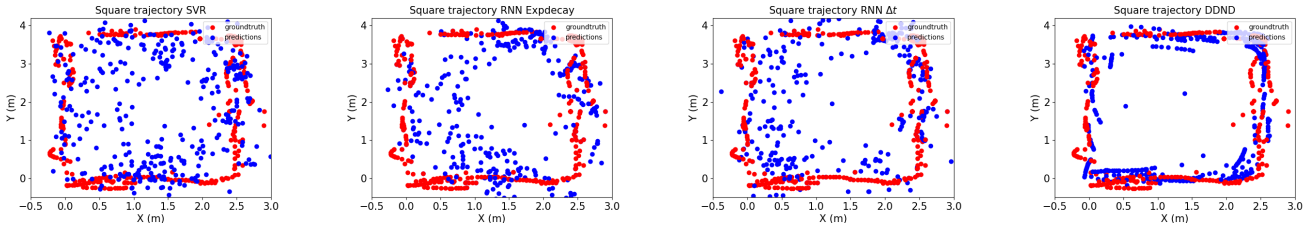
In the following, we present a performance evaluation using real-world mmWave Wi-Fi data from COTS 802.11ad devices.

4.1. mmWave Wi-Fi Localization Dataset

We use 802.11ad-compliant TP-Link Talon AD7200 routers to collect beam SNRs at 60 GHz [30]. A small in-house mmWave Wi-Fi dataset for moving object localization is collected with one router placed on a fixed stand and the other on a TurtleBot as the moving object. The TurtleBot is equipped with a LiDAR and wheel encoder for mapping and gathering location labels. For each responder channel time t_i, the Talon AD7200 router uses 36 directional beampatterns to train the beam, and, hence, we have $M = 36$ for $\mathbf{b}$.

4.2. Implementation

We use a sequence time window of $\Delta T_w = 5$ seconds to group the raw mmWave Wi-Fi beam SNR dataset into sequences with varying numbers of samples, due to the irregular sampling of the beam SNR. We split the sequences into training, validation, and test sets, respectively, with a ratio of $[0.8, 0.1, 0.1]$. We also standardize each entry b_m in the beam SNR by subtracting the mean and normalizing it with the standard deviation. The time vector $\{t_i\}_{i=0}^N$ within each ΔT_w is normalized into $[0, 1]$. We implement the proposed method in the Pytorch framework with a hidden state dimension of $E = 20$ and a latent dimension of $L = 20$. We use the 5th-order Runge-Kutta ODE solver for the decoder. We train the network using the Adamax optimizer with a learning rate of 0.01 and no weight decay. The model is trained with a mini-batch size of 32 sequences, and the loss weighting terms are $\alpha = 0.5$ and $\beta = 0.1$.



(a) (b) (c) (d)

Fig. 3. Visualization of trajectory estimation over selected test sequences: (a) SVR (b) RNN Expdecay (c) RNN Δt (d) DDND

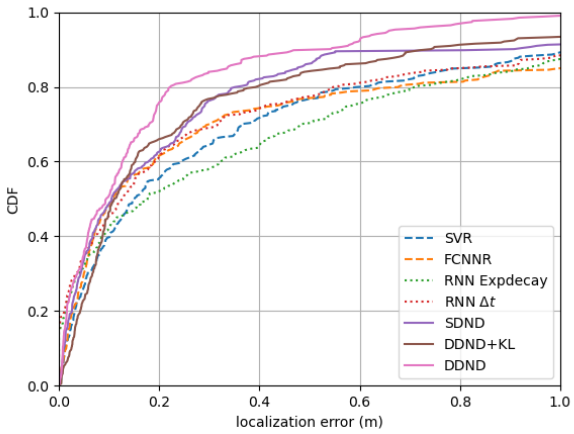


Fig. 4. Cumulative distribution function (CDF) of localization errors.

4.3. Comparison to Baseline Methods

We evaluate the following baseline methods and variants of our framework for the ablation study:

- Frame-based: 1) support vector regressor (SVR); 2) A fully connected neural network regressor (FCNNR);
- Sequence-based: 3) An RNN with exponential decay; 4) RNN with Δt_i attached to the beam SNR;
- Variants of the proposed framework: 5) single-decoder neural dynamic (SDND); 6) DDND with KL divergence (DDND+KL); 7) DDND without KL divergence (DDND).

For the frame-based methods (i.e., 1) and 2)), the coordinate is estimated only from the current frame. The sequenced-based baseline methods (i.e., 3) and 4)) are standard RNNs with the following modifications to handle the irregularly sampled beam SNRs. For the RNN with exponential decay, it applies an exponential decay from one hidden state to the next $\mathbf{h}_i' = e^{-\Delta t_i}\mathbf{h}_i$. For the RNN Δt, the input is the concatenation of the beam SNR $\mathbf{b}_i$ and the sample interval Δt_i. For 5), we first design a single head decoder model without the waveform reconstruction $\mathcal{B}_{\theta_d}(.)$ decoder, while, for 6), we include an additional KL divergence term to our loss function of (10). The localization error is computed as the root mean-squared errors (RMSE) between the estimated $\hat{\mathbf{c}}$ and the ground truth $\mathbf{c}$.

Fig.3 shows the estimated trajectories over test sequences for selected methods. For the frame-based SVR method, the point-like coordinate estimates are scattered within the square trajectory in Fig.3 (a). This is slightly improved by the sequence-based RNN expdecay method in Fig.3 (b) and RNN Δt in Fig.3 (c) as more trajectory estimates are pushing towards the square trajectory. By comparing Fig.3 (d) to Fig.3 (a)-(c), it is clear to see that the proposed DDND is able to learn the underlying dynamics and have more clustered trajectory estimates around the true square trajectory.

Table 1. Localization errors (m) on beamSNR localization dataset.

	Mean	Median	CDF@0.9
SVR	0.42	0.15	1.05
FCNNR	0.46	0.11	1.43
RNN Expdecay	0.40	0.18	1.12
RNN Δt	0.33	0.13	1.09
SDND (ours)	0.34	0.11	0.88
DDND+KL (ours)	0.26	0.11	0.74
DDND (ours)	**0.17**	**0.09**	**0.52**

Fig.4 compares the baseline methods against the proposed solution in terms of the cumulative distribution function (CDF) of the localization error. It shows that proper learning and modeling of the latent dynamics and trajectory motion of the moving object from irregularly-sampled Wi-Fi data can improve the localization performance over the frame-based and traditional sequence-based methods. The proposed method and its variants (5), 6) and 6)) are seen to have significantly fewer large localization errors in Fig.4. For instance, the proposed DDND method has 10% fewer localization errors that are larger than 0.6 m than all baseline methods.

Table 1 further summarizes quantitative performance in terms of the mean, median, and location error corresponding to the 90th percentile of the CDF. It further confirms that the proposed DDND can better deal with the irregularly sampled beam SNRs and model the underlying dynamics of the latent space in a continuous-time fashion. Compared with the modified RNN methods, the DDND method almost reduces the localization error by half for all three performance metrics.

5. CONCLUSION

This paper tackles the problem of intermittently-sampled mmWave Wi-Fi beam training measurements for localization. Specifically, we proposed the dual-decoder neural dynamic framework that learns the continuous inherent latent dynamics. Performance comparison confirms the performance gain of the proposed method. We plan to scale up the mmWave Wi-Fi dataset by including localization scenarios over multiple trajectories, considering the usage of multiple devices as well as the fusion of multi-band measurement features.

6. REFERENCES

[1] P. Bahl and V. N. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," in *INFOCOM*, March 2000, vol. 2, pp. 775–784.

[2] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Comput. Surv.*, vol. 46, no. 2, Dec. 2013.

[3] S. He and S. H. Chan, "Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons," *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 466–490, 2016.

[4] K. Wu et. al., "CSI-based indoor localization," *IEEE Trans. on Parallel and Distributed Systems*, vol. 24, no. 7, pp. 1300–1309, July 2013.

[5] M. Pajovic, P. Wang, T. Koike-Akino, H. Sun, and P. V. Orlik, "Fingerprinting-based indoor localization with commercial MMWave WiFi–Part I: RSS and Beam Indices," in *GLOBECOM*, Dec. 2019.

[6] P. Wang, M. Pajovic, T. Koike-Akino, H. Sun, and P.V. Orlik, "Fingerprinting-based indoor localization with commercial mmwave WiFi–Part II: Spatial beam SNRs," in *GLOBECOM*, Dec 2019.

[7] T. Koike-Akino, P. Wang, M. Pajovic, H. Sun, and P. V. Orlik, "Fingerprinting-based indoor localization with commercial mmwave WiFi: A deep learning approach," *IEEE Access*, vol. 8, pp. 84879–84892, 2020.

[8] P. Wang, T. Koike-Akino, and P. V. Orlik, "Fingerprinting-based indoor localization with commercial mmwave WiFi: NLOS propagation," in *GLOBECOM*, December 2020.

[9] J. Yu, P. Wang, T. Koike-Akino, and P. V. Orlik, "Human pose and seat occupancy classification with commercial mmwave WiFi," in *GLOBECOM Workshop on Integrated Sensing and Communication (ISAC)*, December 2020.

[10] J. Yu, P. Wang, T. Koike-Akino, Y. Wang, P. V. Orlik, and R. M. Buehrer, "Multi-band Wi-Fi sensing with matched feature granularity," *IEEE Internet of Things Journal*, pp. 1–1, 2022.

[11] T. Koike-Akino, P. Wang, and Y. Wang, "Quantum transfer learning for Wi-Fi sensing," in *IEEE International Conference on Communications (ICC)*, May 2022.

[12] 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, Jan 2017.

[13] H. Chen, Y. Zhang, W. Li, X. Tao, and P. Zhang, "ConFi: Convolutional neural networks based indoor Wi-Fi localization using channel state information," *IEEE Access*, vol. 5, pp. 18066–18074, 2017.

[14] H. Xia, P. Wang, T. Koike-Akino, Y. Wang, P. V. Orlik, and Z. Ding, "Adversarial bi-regressor network for domain adaptive regression," in *International Joint Conference on Artificial Intelligence (IJCAI)*, 2022, pp. 3608–3614.

[15] J. Yu, P. Wang, T. Koike-Akino, and P. V. Orlik, "Multi-modal recurrent fusion for indoor localization," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2022.

[16] M. Youssef and A. Agrawala, "The horus location determination system," *Wirel. Netw.*, vol. 14, no. 3, pp. 357–374, June 2008.

[17] S. Mazuelas et. al., "Robust indoor positioning provided by real-time RSSI values in unmodified WLAN networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 5, pp. 821–831, Oct 2009.

[18] X. Wang, L. Gao, and S. Mao, "CSI phase fingerprinting for indoor localization with a deep learning approach," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 1113–1123, Dec 2016.

[19] X. Wang, L. Gao, and S. Mao, "Biloc: Bi-modal deep learning for indoor localization with commodity 5GHz WiFi," *IEEE Access*, vol. 5, pp. 4209–4220, 2017.

[20] C. Hsieh, J. Chen, and B. Nien, "Deep learning-based indoor localization using received signal strength and channel state information," *IEEE Access*, vol. 7, pp. 33256–33267, 2019.

[21] M. T. Hoang, B. Yuen, X. Dong, T. Lu, R. Westendorp, and K. Reddy, "Recurrent neural networks for accurate RSSI indoor localization," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10639–10651, Dec 2019.

[22] X. Wang, L. Gao, S. Mao, and S. Pandey, "DeepFi: Deep learning for indoor fingerprinting using channel state information," in *WCNC*, March 2015, pp. 1666–1671.

[23] A. S. Paul and E. A. Wan, "Wi-Fi based indoor localization and tracking using sigma-point Kalman filtering methods," in *PLANS*, 2008, pp. 646–659.

[24] J. Wang and J. G. Park, "An enhanced indoor ranging method using CSI measurements with extended Kalman filter," in *PLANS*, 2020, pp. 1342–1348.

[25] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 11 1997.

[26] J. Ding and Y. Wang, "WiFi CSI-based human activity recognition using deep recurrent neural network," *IEEE Access*, vol. 7, pp. 174257–174269, 2019.

[27] Y. Rubanova, R. Chen, and D. K. Duvenaud, "Latent ordinary differential equations for irregularly-sampled time series," *Advances in neural information processing systems*, vol. 32, 2019.

[28] R. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, "Neural ordinary differential equations," *Advances in neural information processing systems*, vol. 31, 2018.

[29] W. Cao, D. Wang, J. Li, H. Zhou, L. Li, and Y. Li, "Brits: Bidirectional recurrent imputation for time series," *Advances in neural information processing systems*, vol. 31, 2018.

[30] D. Steinmetzer, D. Wegemer, and M. Hollick, "Talon tools: The framework for practical IEEE 802.11ad research," in *Available: https://seemoo.de/talon-tools/*, 2018.