University of Stuttgart
Institute of Telecommunications

On Deep Learning-based Massive MIMO Indoor User Localization

M. Arnold, S. Dörner, S. Cammerer, S. ten Brink
University of Stuttgart, Institute of Telecommunications, Pfaffenwaldring 47, 70569 Stuttgart, Germany
{arnold,doerner,cammerer,tenbrink}@inue.uni-stuttgart.de

ABSTRACT

We examine the usability of deep neural networks (NNs) for multiple-input multiple-output (MIMO) user positioning solely based on the orthogonal frequency division multiplex (OFDM) complex channel coefficients:

- Deployed on top of an existing OFDM MIMO system
- → Does not require any additional piloting overhead
- Line of sight (LoS) and non-line of sight (NLoS) measurements provided
- Neural network training by stochastic gradient descent (SGD)
  - – Requires a large amount of data-points for training
  - → We propose a two-step training procedure
- Two-step training of the NN:
  - – First step: Extensive training on simulated line of sight (LoS) data
  - – Second step: Finetuning training on measured NLoS data
  - → Reduces the amount of required training positions
  - → Reduces the effort for data acquisition

The need for indoor positioning systems (IPSs)

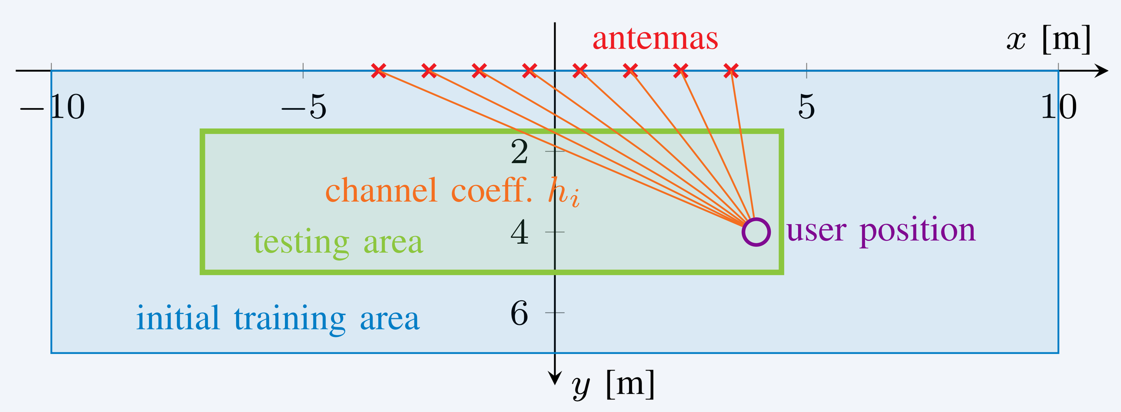
- Enabler for a wide range of applications e.g. navigation, smart factories, Internet of Things (IoT) network sensors
- Improvements of communications algorithms like beamforming or channel estimation based on motion prediction
- Outdoor positioning sufficiently solved by satellite systems
- Indoor positioning systems are diverse and highly application optimized
- Current approaches can be split coarsely into two categories:
  1. **Model-based:** Position estimation based on how the channel is expected to behave
  2. **Data-driven:** Interpolation in-between collected features (often called fingerprints) stored in a database

Why based on OFDM MIMO systems?

- OFDM is the workhorse of many state-of-the-art standards
- → Widely used among mobile communication devices
- MIMO systems provide detailed channel characteristics due to multiple antennas
- → Already used by many recent devices, likely to increase in future
- → **No additional piloting required!**
  We simply use already available channel coefficients

Background

Basic System Model



- Line array positioned next to training and testing area
- User (single antenna) positioned within training or testing area
- Each antenna has a slightly different channel observation to the user due to different spatial antenna positions
- → OFDM (single tap eq.) channel coefficients h_i differ per antenna

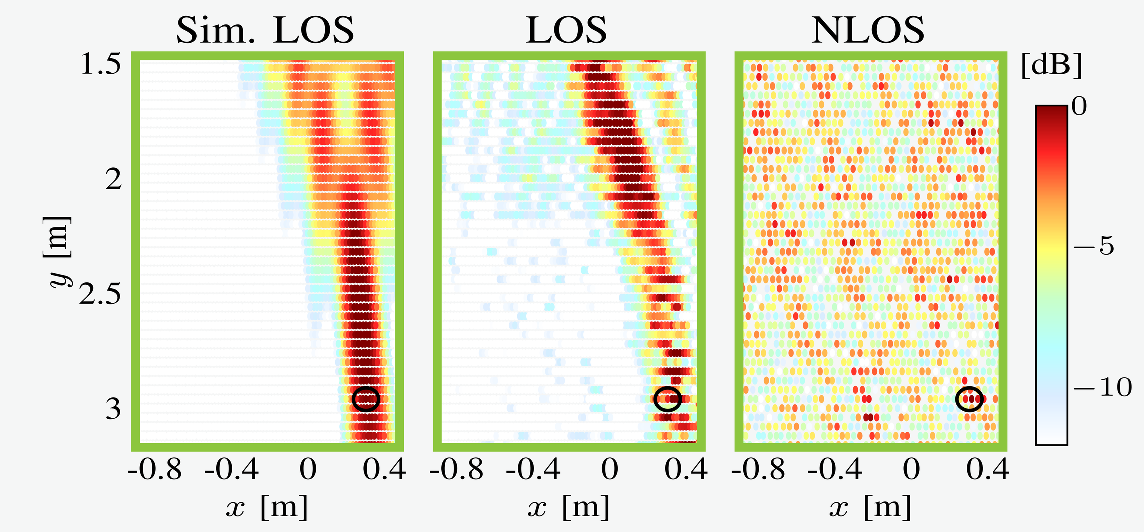
Two step training process

1. **Initial training** on simulated LoS channel coefficients out of training area (blue)
   - Channel coefficients can be computed for all user positions
   - Unlimited amount of training data for arbitrary large area
   - Eliminates overfitting since no data-point is used twice
   - → Results in a better weight initialization for final training
2. **Finetuning training** on measured channel coefficients out of the testing area (green)
   - Final testing area lies within pre-trained area
   - Limited amount of measured training data available
   - → Fewer data-points and faster convergence is key

References

[1] M. Arnold, M. Gauger, and S. ten Brink, "Evaluating massive MIMO precoding based on 3D-channel measurements with a spider antenna," in ISWCS, Aug 2017, pp. 134–139.

Simulated and Measured Data



This figure shows the spatial energy map over the testing area for maximum ratio (MR) precoding for a target user (black circle). It shall illustrate the spatial conditions for different channel setups.

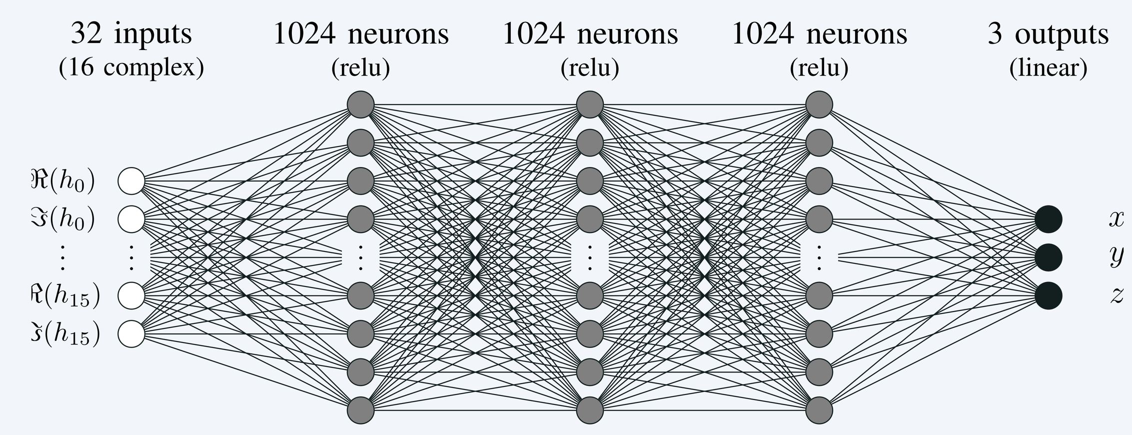
Simulated Channels:

$$h_{i,\text{LoS}} = \left(\frac{\lambda}{4\pi d}\right) e^{j2\pi\frac{d}{\lambda}} \qquad (1)$$

Measured Channels:

- Spatialy consistent channel measurements conducted in [1]
- Setup: 16 antennas, line array, $\lambda/2$ distance, $f_c = 2.35$ GHz
- Position labels obtained by spider antenna stepping motors [1]
- NLoS environment enforced by a metal plate in line of sight
- About $60,000$ data-points for LoS and NLoS measurement each
- Measured area size: 1.35m x 1.78m, distance to line array: 1.48m

Deep Neural Network Structure



NN Parameters:

- Simple dense feed-forward layers, $2,136,067$ weights in total
- Optimization on mean squared error (MSE) with SGD (Adam)
- Activation: rectified linear unit (ReLU) $g_{\text{ReLU}}(x) = \max\{0, x\}$

Performance Metric: normalized mean squared error (NMSE)

MSE normalization with respect to distance $d = \|\mathbf{p}\|^2$:

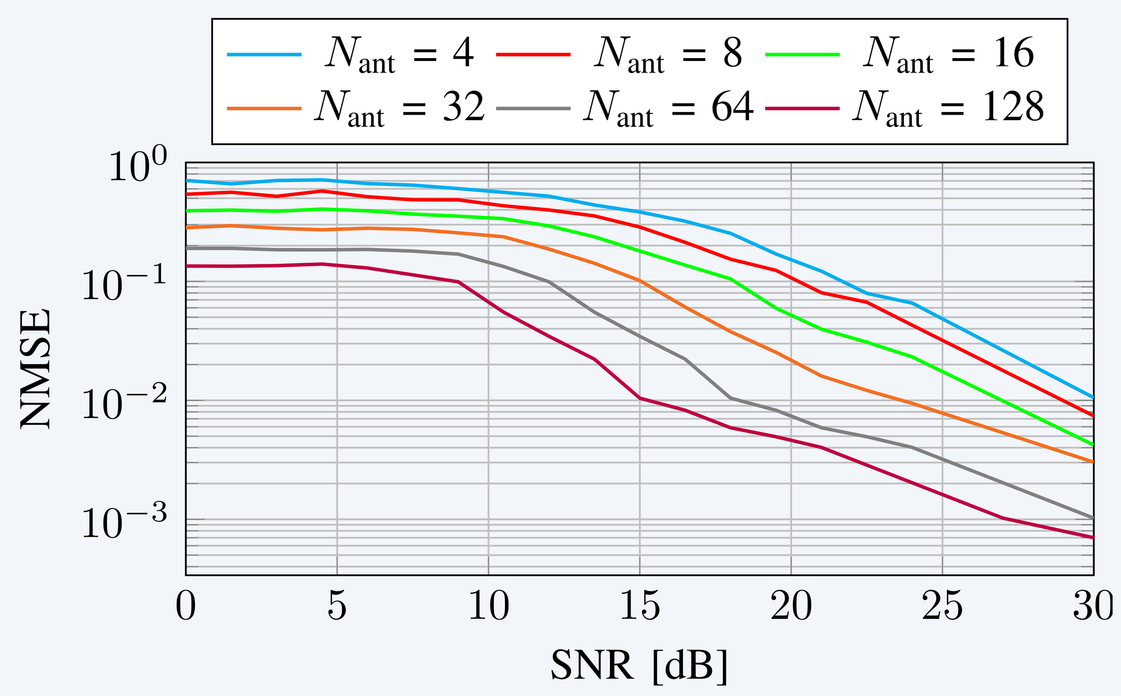
$$\text{NMSE} = \mathbb{E}\left[\frac{\|\mathbf{p} - \hat{\mathbf{p}}\|^2}{\|\mathbf{p}\|^2}\right] \qquad (2)$$

Where $\mathbf{p}$ is the actual position and $\hat{\mathbf{p}}$ the estimated position

- A user position close to the base station is easier to estimate than a position with a large distance
- 1% NMSE means e.g. 1cm error for a position that is 1m away
- → Normalization simplifies comparison of different scenarios

Initial Training on Simulated LoS Channels

NMSE Performance on Simulated Channel



This figure shows the NNs NMSE performance over SNR and varying number of antennas N_{ant} on the simulated LoS model.

Observations:

- → $N_{\text{ant}} = 4$ at SNR of 30dB gives an accuracy of 1% NMSE
- → Doubling N_{ant} leads to a ≈3dB gain in NMSE, as expected

Definition of the signal-to-noise-ratio (SNR):

$$\text{SNR} = \frac{\sum_{n=1}^{N_{\text{ant}}} |h_{\text{LoS}}(n)|^2}{\sigma^2} \qquad (3)$$

Finetuning on Measured LoS/NLoS Data

Data-point Distribution

General observation:

The NN does not perform well outside of the trained area!

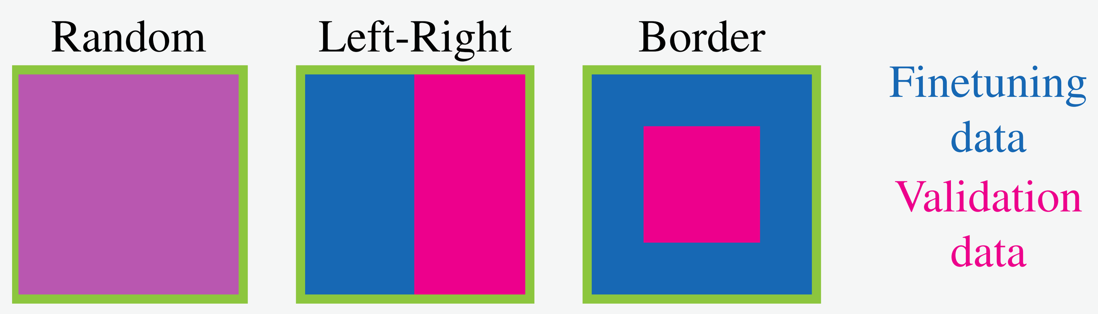
→ It seems to learn fingerprints, but NOT a global solution

This raises two questions:

1. How many data-points are needed for sufficient training?
2. Where do these measurements have to be located?

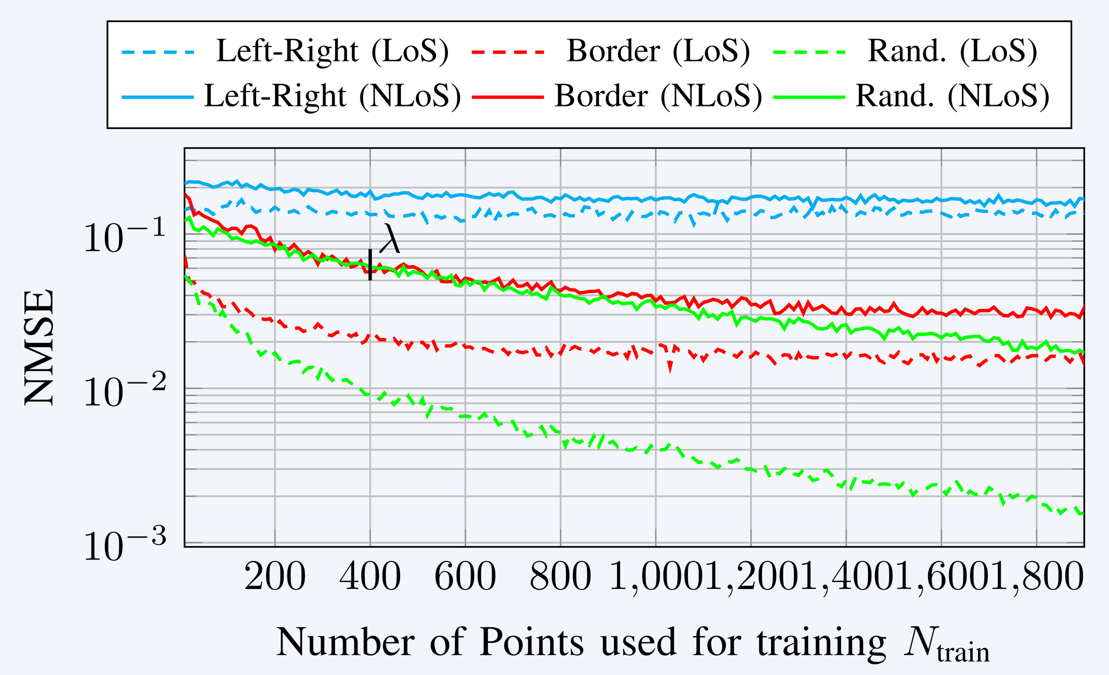
To challenge the second question we came up with

Three scenarios to select finetuning training data-points:



1. **Random:** Pick data-points randomly out of the whole area
2. **Left-Right:** Only use data-points out of one side of the area
3. **Border:** Only use data-points out of the border area (30%)

NMSE Performance over number of points used

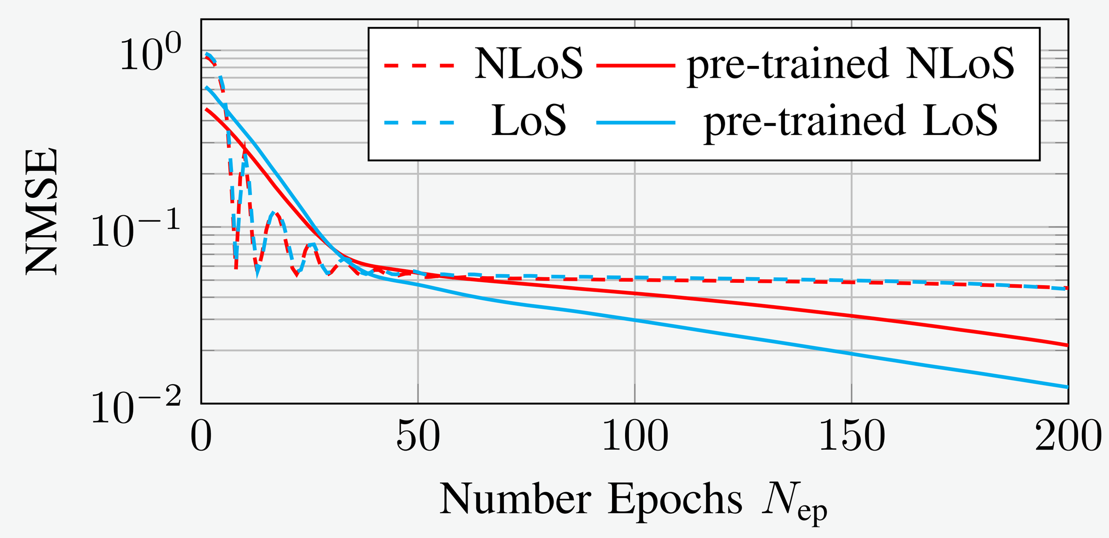


This figure shows the NNs NMSE performance after finetuning on measured LoS (solid) and NLoS (dashed) data for all three scenarios over the amount of used finetuning data-points. $N_{\text{ep}} = 100$ epochs are used for training.

Observations:

- The random scenario always performs best
- The left-right scenario does not converge, stops at ≈11% NMSE
- The border scenario converges to a reasonably good accuracy
- About 400 points are sufficient for the border case to achieve a reasonably well accuracy → accuracy of λ sampling achieved
- The NLoS cases generally show a lower accuracy, presumably due to the more complex channel characteristics
- In the NLoS case the border scenario performs as good as in the random scenario up to 1000 data-points

NMSE Performance over number of training epochs



This figure shows the NMSE performance of a pre-trained NN and a randomly initialized NN trained on 200 measured data-points for both the LoS and the NLoS case.

Observations:

- The pre-trained NN reaches a lower final NMSE accuracy
- The uninitialized NN needs more data-points for equal accuracy

Conclusion

- NNs can be used for user localization in MIMO OFDM systems
- Pre-training the NN with a simulated channel model
  - – reduces the amount of data-points needed for training
  - – leads to a higher final accuracy
  - – leads to a faster convergence in terms of training time
- An accuracy of less than 1% NMSE can be reached for this area