USC Viterbi
School of Engineering

Ming Hsieh
Department of Electrical Engineering

ITU Workshop, San Jose, CA, USA, August 2018

Machine learning for wireless propagation channels

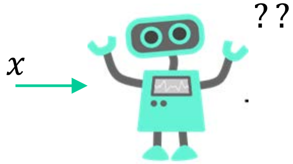
Andreas F. Molisch
Wireless Devices and Systems (WiDeS) Group
Viterbi School of Engineering
University of Southern California (USC)

USC

Table of contents

Machine Learning

- Machine Learning (ML) allows the computer to use set of observations to perform certain tasks.



- **Components in ML:**
  - Learning objective.
  - Data.
  - Model and learning.

- **Types of ML:**
  - Unsupervised Learning. (e.g., clustering)
  - Supervised Learning (classification and regression)
  - Reinforcement learning. (learning policy with trail and error)
  - Others (online learning, semi supervised .)

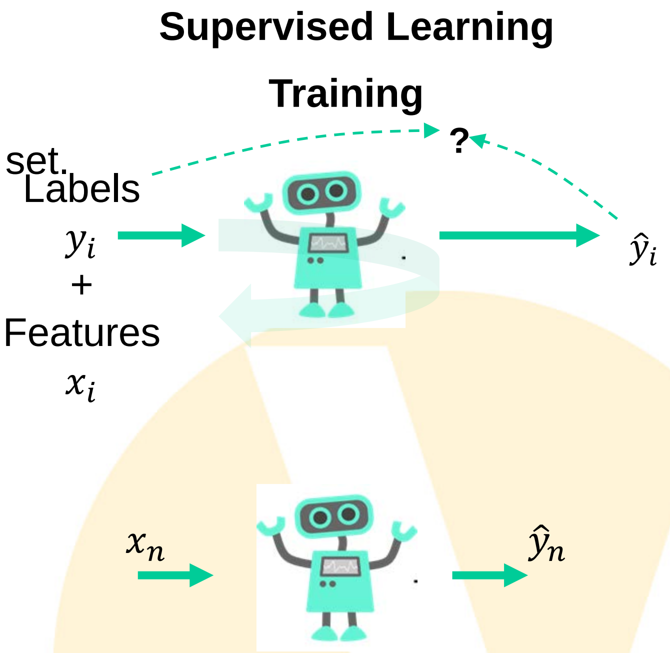
The application and the scenario identifies the type of ML

Supervised ML

- Available labeled data can be used to *train* the ML model.

- **Problem types:**
  - Classification
    e.g., labels (aka classes) belong to finite countable set.
  - Regression,
    Labels could be continuous real line.

- **Models include:**
  - Linear regression.
  - Logistic regression.
  - Decision trees.
  - Neural Networks.
  - Support vector machines.



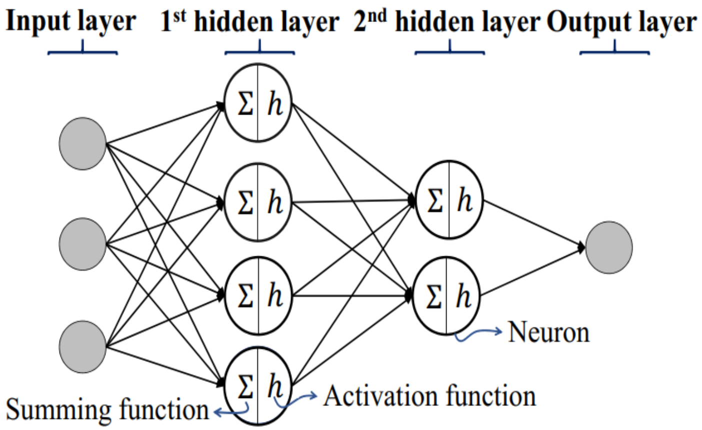
Applications in Wireless Communication

- ML can be used to perform complicated task where mathematical models are very complex or difficult to capture.
- There have been large literature:
  - Different wireless technologies such as wireless sensor, ad hoc and cellular networks.
  - Centralized and distributed systems.
  - All network layers.
- Example of applications in *Physical* layer:
  - Ranging and localization, e.g., [1].
  - Performance prediction and channel Estimation, e.g.,[2,3].
  - Resource allocation, e.g.,[4].
  - Beam prediction, e.g., [5].
  - Modulation recognition[8].

Neural Networks

- Computational model that mimic the neurons in the brain.
- Multi layer Neural Network (NN) was proven to be *universal* function approximator [6].
- Proposed to model many aspects in wireless communication, such as [7]:
  - Amplifiers' non linearity.
  - Equalization.
  - Decoding in spread spectrum.
- Used also as channel model and prediction [10,11].
- Backpropagation is a popular method that use gradients to train NN.

Feedforward NN
(multilayer perceptron)

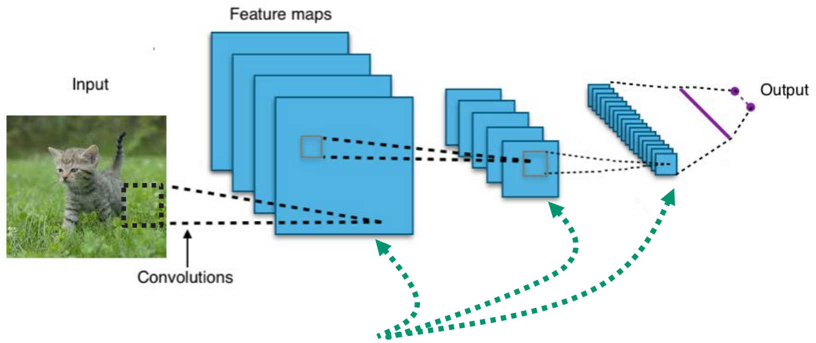


Deep Neural Networks

- Is *hierarchical* learning.
- *Most m*odern models are based on *multi layer NNs*.
- Although ML and NN existed long before today, deep learning became relevant today due [9]:
  - The remarkable success in language and image recognition and computer vision.
  - The advancement in hardware and parallel computing capabilities.
  - Availability of many deep learning open source software.
  - The big data era! which challenges analysis with conventional mathematical models.
- DL has also found its way to wireless communication with the emerge of new complex channels and systems, for instance:
  - Massive MIMO.
  - Multi band communication (mmWave and cmWave).
  - Underwater and molecular channels.
  - Cognitive radio systems.

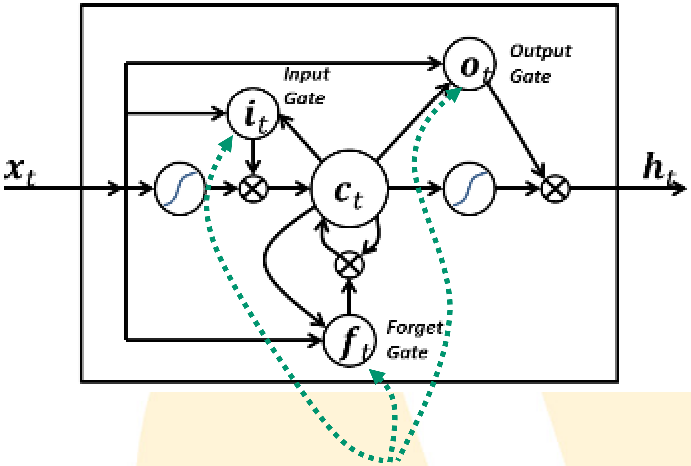
Example of Deep Neural Networks Blocks

Convolutional NN (CNN)
(popular in computer vision)



Filters that extract features and do .

Long Short-Term Memory (LSTM)
(Popular in text prediction, speech recognition, Natural language processing)



Is a *Recurrent* NN (RNN), it uses gates that control what data to be remembered, released and erased

Table of contents

Multi Layer Neural Networks for Channel Prediction

- Several empirical models exist to predict the pathloss value, eg., Okumura-Hata (OH) model and Ericsson 9999 model.
- Other deterministic models that depends, for instance, on topographic database can be used.
- Empirical Models are simple to use but lack accuracy, while deterministic models are complex and provide better prediction.
- NN can be a solution to the trade-off, as it is simple to use and can capture complicated environment efficiently.

Multi Layer NN for Channel Prediction: Example

- In [25], the authors use different NN architectures to predict the path loss in a macrocell for a rural area.
- The feature input are:
  - Tx-Rx Distance, BS antenna height, terrain clearance angle (TCA), terrain usage, vegetation type (VT) and vegetation density around rx.
- Training Testing Scenarios:
  - scenario-1: same cell different route: 4 routes for training 1 route for testing.
  - scenario-2: Different cells: train one cell and test on another.
  - scenario-3: All cells: randomly chosen points.
- Example for an NN with (6,3,1), the average error $\bar{\epsilon}$, and standard deviation σ_{ϵ}
- For scenario-1: $\bar{\epsilon} = -3.5$ dB with $\sigma_{\epsilon} = 7.4$ dB, compared to $\bar{\epsilon} =$-12.8 and $\sigma_{\epsilon} = 8.7$ dB for OH.

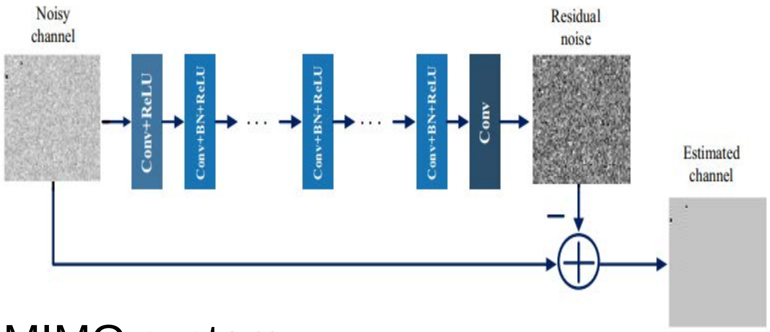
For scenario-2: $\bar{\epsilon} = -4.9$ dB with $\sigma_{\epsilon} = 8.3$ dB, compared to $\bar{\epsilon} =$-12.4 and $\sigma_{\epsilon} = 11.4$ dB for OH.

Channel Models with multilayer NN

- For indoor [24], at 900 MHz, use position, antenna heights, gains, type of indoor room, and some other penetration information of the direct path, eg., number of walls in the direct path, error is reported as an RMSE of 4.4 dB.
- [11] Similarly for indoor, with slightly different feature details, such as the use of the percentage of walls, rooms etc., train in one building and test in another one. Report error around 2 dB with standard deviation 7 dB.
- For macrocells, [26] use databases for land use/land cover information, and with error ~ 0, and standard deviation ~ 8dB.
- [10] Use 27 input features that describe the several rays, which are derived mainly from geometrical of the environment and the tx-rx locations, gives an average error of ~ 1 dB.
- [27] Propose hybrid model that combine the empirical models (eg, OH) with NN, in suburban environment, show good prediction capability.

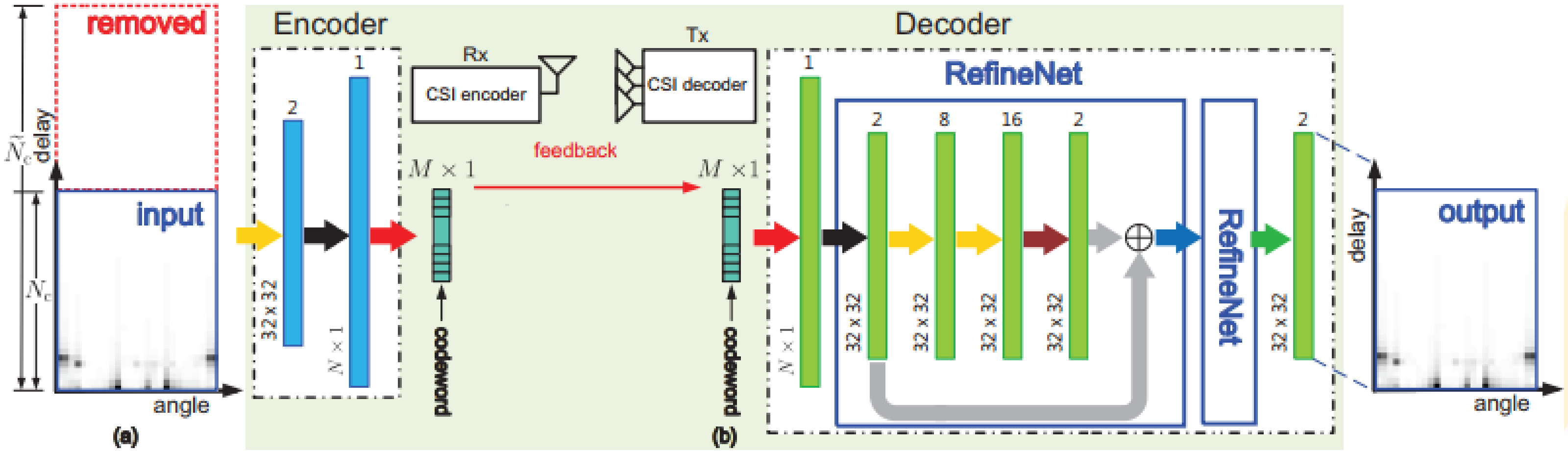
Application in Wireless-Channel Estimation in Massive MIMO

- [20] views channel matrix as 2D natural image.
- apply approximate massage passing NN that is based on denoising convolutional neural network (DnCNN).



- [21] aim sto reduce the CSI feedback in massive MIMO system.

The proposed Deep Network learns a transformation from CSI to a near-optimal number of representations and an inverse transformation from codewords to CSI.



Application in Wireless – Others

- [13] uses Deep Learning (LSTM and CNN layers) for *detection* molecular channel where no knowledge of the channel is assumed.
- *Detection* over conventional channels discussed, e.g., for OFDM system [12] and MIMO system [15].
- In general there have been number of studies that *unfold iterative schemes* through multi-layer NN, eg., for decoding using belief propagation [18].
- Several other works used multi-layer NN to predict channel parameters, such as AoA at the device side using the observations at the BS [19].

Table of contents

- Background and motivation
- Survey: for what channel-related probems can machine learning be used?
- Example: switching in dual-band (cm/mm-wave) systems
- Example: remote CSI inference
- Summary and conclusions

Dual- (Multi-) Band Systems

- Promises higher throughput and quality of service.
- Different structure, e.g. :
  - cmWave used for control and mmWave for data.
  - Joint data transmission over cmWave and mmWave.
  - cmWave as backup for data transmission.
- Why Deep Learning?
  - Dual Band is unconventional channel, with complex join propagation properties.
  - Base station have access to data.
  - Pilot training on both bands is expensive.

Learning In Stochastic Environment [22]

- Given "features", how to the Band Assignment (BA)?
- Model the band assignment as a *classification* problem, "1" assign to mmWave band.
- Stochastic channel based on a proposed jointly normal power distribution [16].
- Proposed a Threshold Based Band Assignment (TBBA) solution using the joint normal power distribution.
- TBBA requires the knowledge of distance and channel statistics.
- Linear and Logistic Regression (LR and GR) in addition to multi layer NN

> Learning based (especially multi layer NN) could outperform the theoretical scheme.

Variable	Band c/m
f_b	2.5/28 GHz
Bandwidth ω_b	10/100 MHz
P^b_{tx}	15/15 dBm
ϵ	4
d_{break}	30 m
d_{dcor}	18/13
σ_b	5/7 dB
$\rho_{m,c}$	0.75
Noise Spectral Density	-174 dBm/Hz

TABLE I
STOCHASTIC CHANNEL SIMULATION CONFIGURATIONS

Feature / Combination	c-1	c-2	c-3	c-4	c-5
d	✓	✓	✓	✓	
θ	✓	✓			
cmWave Power		✓	✓		✓
NN $\bar{\mathcal{E}}_S$	.15	.119	.118	.156	$\boxed{.119}$
GR $\bar{\mathcal{E}}_S$	.155	.119	.119	.155	$\boxed{.12}$
LR $\bar{\mathcal{E}}_S$	.155	.122	.122	.155	$\boxed{.119}$

TBBA / γ_T choice	$\gamma_T = .5$	γ_T function of distance
TBBA $\bar{\mathcal{E}}_S$	$\boxed{.18}$	.149

TABLE II
PERFORMANCE OF THE LEARNING OVER THE STOCHASTIC DATA UNDER DIFFERENT FEATURE AVAILABILITY. NOTE THAT ON AVERAGE 24.4% OF THE LABELS ARE "1".

Ray tracing based

- Simulated a more realistic channel through a ray tracer.
- Total of 1184 points in each bands.
- Considered several NN structures.

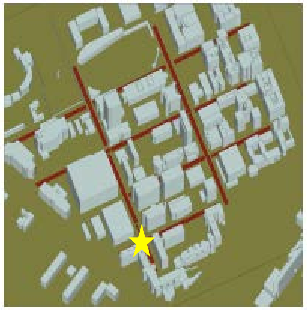


Fig. 2. Ray-tracing simulation environment.

Feature /Combination	c-1	c-2	c-3	c-4	c-5	c-6	c-7	c-8
d	✓	✓	✓			✓		
θ	✓	✓						
cmWave Power		✓	✓	✓	✓		✓	
Delay				✓			✓	✓
AoD							✓	✓
Numb Layers/ α/ γ_L	2/.15/.45	4/.15/.55	1/.05/.5	1/.1/.6	2/.1/.35	3/.5/.45	4/.3/.6	3/.5/.55
NN $\bar{\mathcal{E}}_S$	.078	.061	.072	.074	.085	.182	.067	.093
GR $\bar{\mathcal{E}}_S$	.178	.062	.082	.081	.083	.183	.082	.182
LR $\bar{\mathcal{E}}_S$	.176	.078	.088	.078	.081	.178	.072	.188

TABLE IV
PERFORMANCE OF THE LEARNING TECHNIQUES ON RAY-TRACING DATA, UNDER DIFFERENT FEATURE AVAILABILITY, NOTE THAT THE PERCENTAGE OF POINTS WITH LABELS EQUAL TO "1" IS APPROXIMATELY 30%.

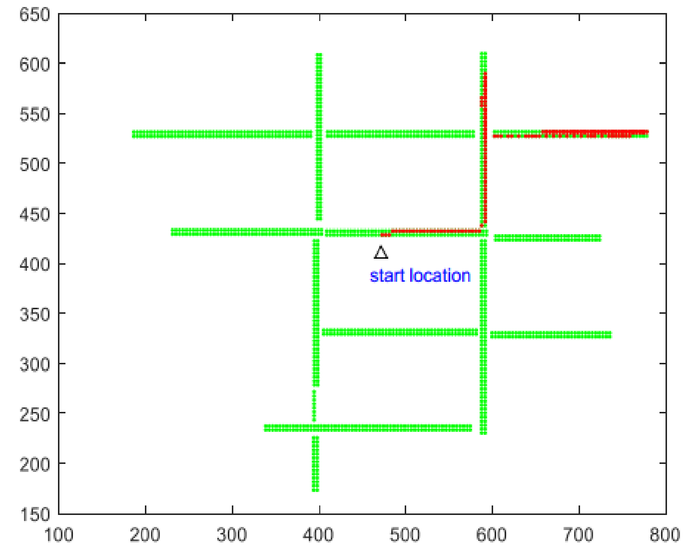
Variable	Band c/m
f_b	2.5/28 GHz
Ant. Pattern	Isotropic
Ant. Polarization	Vertical
P_{tx}^{b}	15/30 dBm
BS height	45 m
MS height	2 m
Max. Diffraction	2/1
Max. Reflection	10

TABLE III
RAY-TRACING SIMULATION CONFIGURATIONS OF USC CAMPUS

- ➢ Multi layer NN show consistent good BA decisions.
- ➢ The use of cmWave power improves the BA accuracy.
- ➢ BA based on distance only has the worse performance.

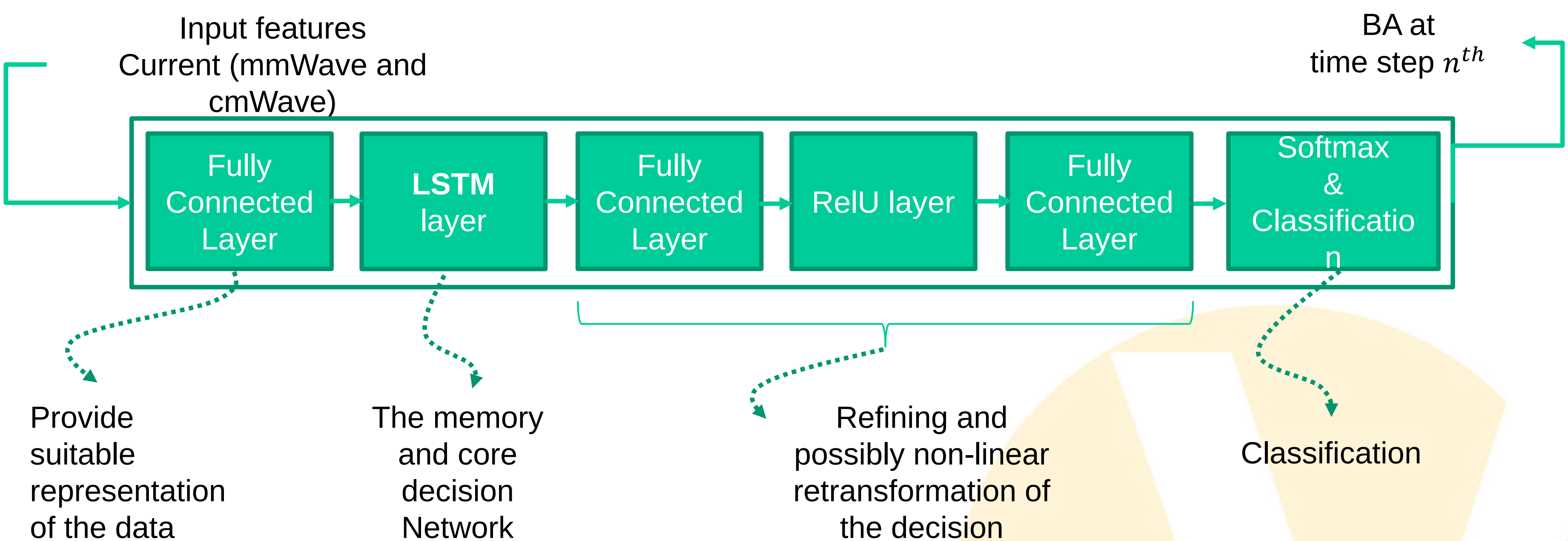
Sequential BA For Mobile Users [23]

- Device move in correlated environment.
- Correlation difficult to capture analytically.
- Given current observed features along with the previous decisions what is the BA for a *future* time step?
- Used 1000 sequences generated over the ray tracing environment.
- Initially assumed that both cmWave and mmWave are observable.



Motion using Semi-Markov Smooth Mobility Model over the grid. Need to predict the BA after n steps

Proposed Network



Result of Sequential Data results

- Used 200 sequences for training 800 for testing.
- Naïve use the current best band as future BA.

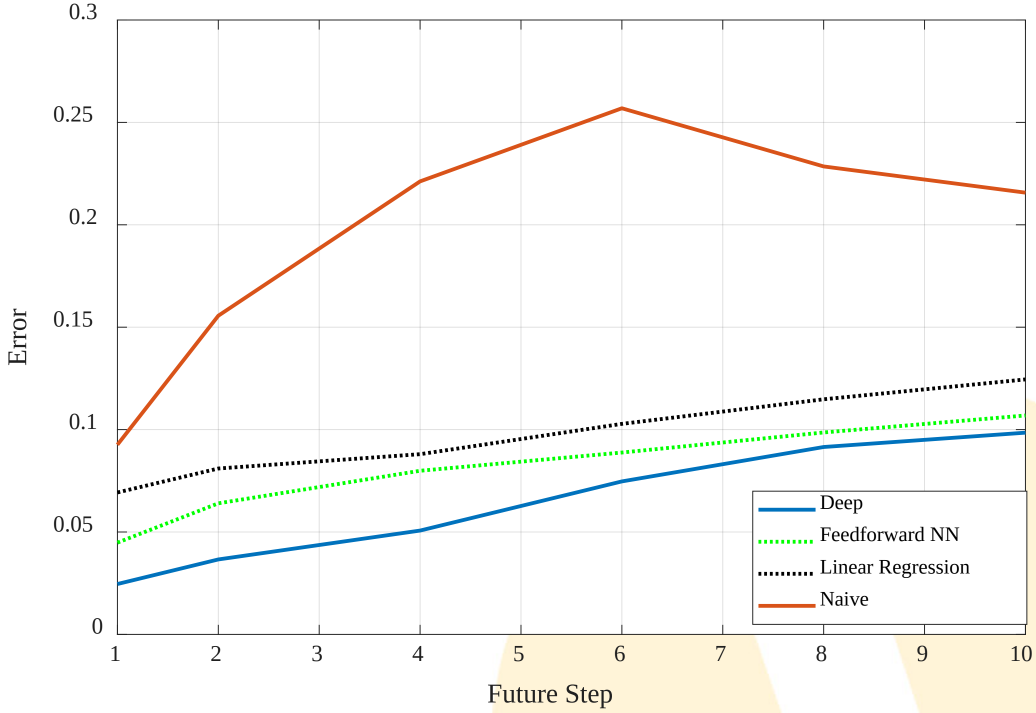


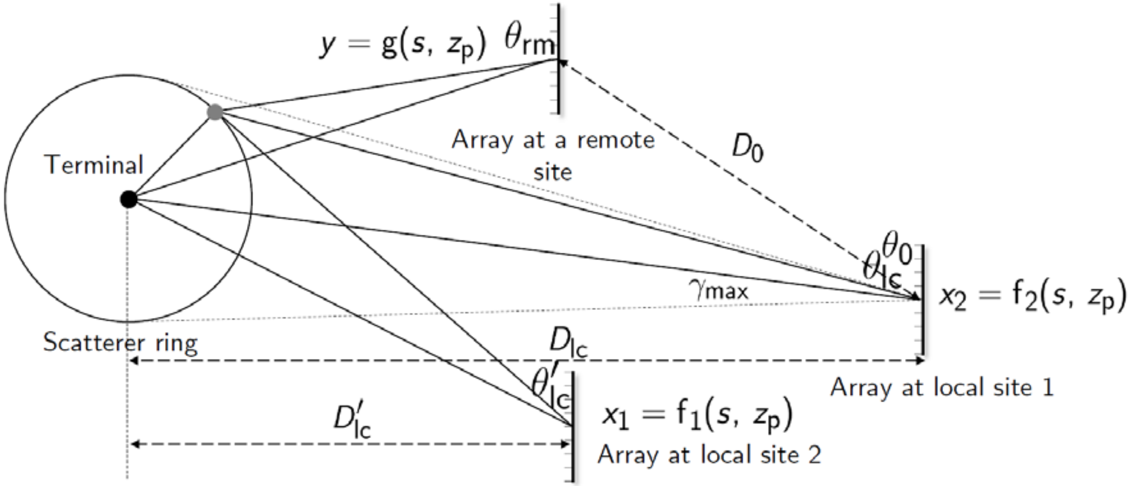
Table of contents

Motivation [30]

- Goal: determine DoA spectrum (or beamforming codebook entry) at remote BS if channel at macro BS is known
- Purpose: reducing overhead of channel estimation

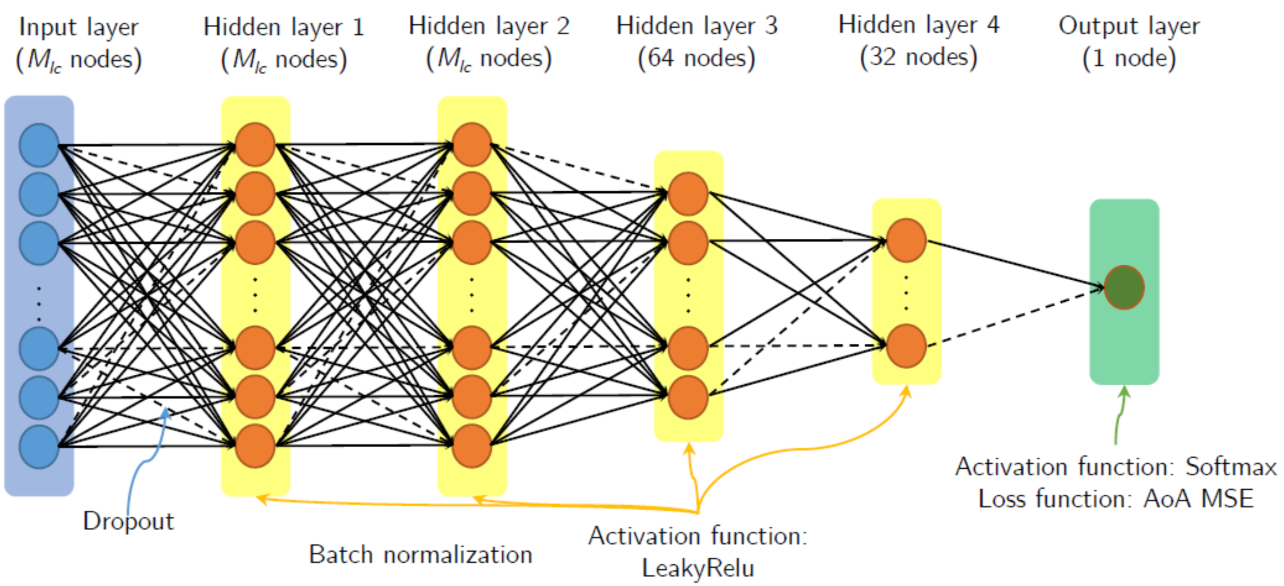
P1: Estimate $\boldsymbol{h}_{\mathrm{rm}}$, given $\boldsymbol{h}_{\mathrm{lc}}$, s.t.,

$$h_{\mathrm{lc},i} = \mathrm{f}\left(\theta_{\mathrm{lc}}, S_{\mathrm{lc}}(\gamma), \phi_{\mathrm{lc},\gamma}, r_{\mathrm{max,lc}}\right),$$
$$h_{\mathrm{rm},i} = \mathrm{g}\left(\theta_{\mathrm{rm}}, S_{\mathrm{rm}}(\gamma), \phi_{\mathrm{rm},\gamma}, r_{\mathrm{max,rm}}\right),$$



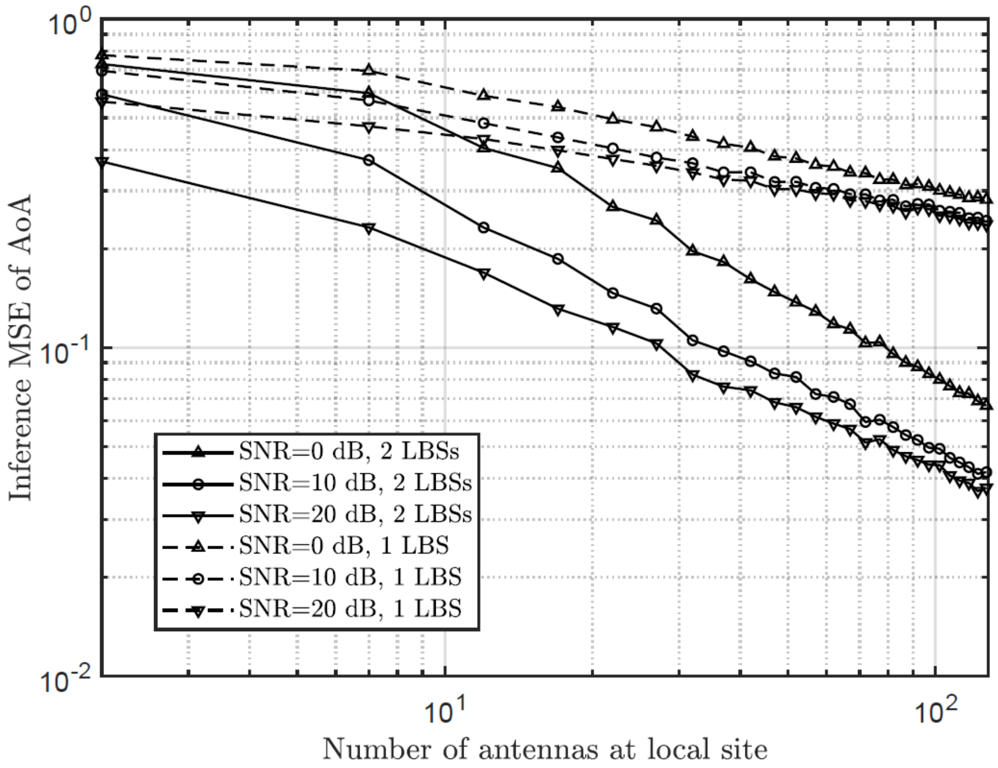
Neural network architecture

- 4 layer structure
- Nonlinear function for hidden layers: LeakyRLu
- Nonlinear function for output sigmoid



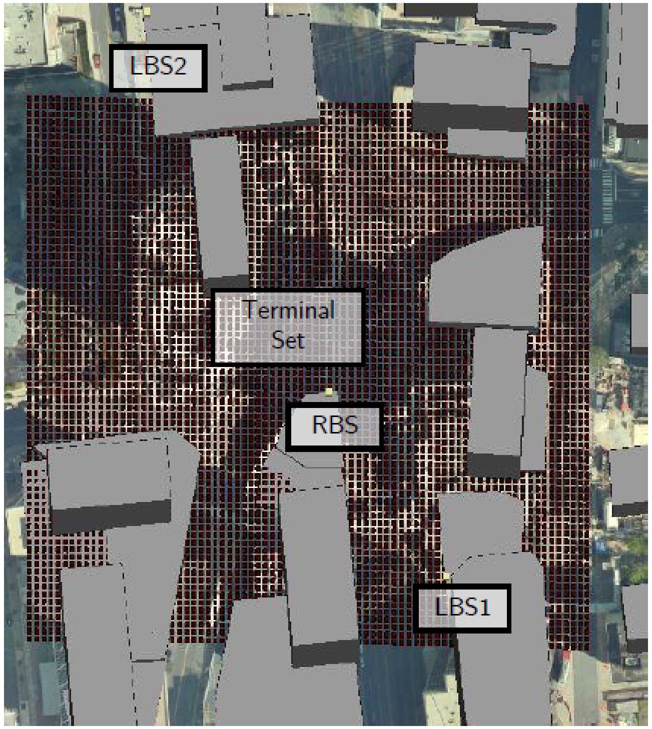
Performance for one-ring model

- Comparable to CRLB



Performance for ray tracing model

- 85% of time error is less than 10%



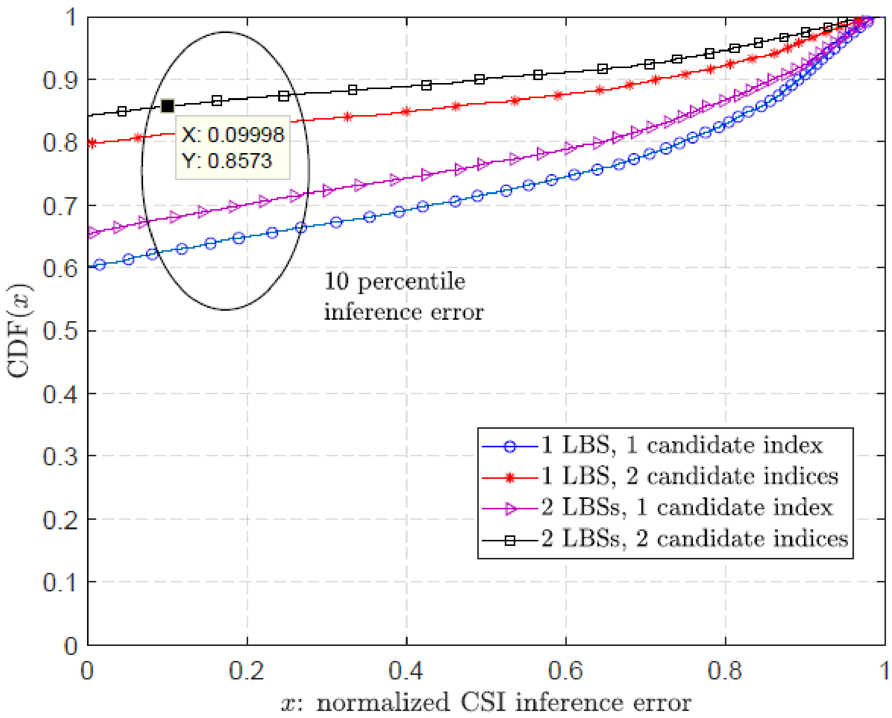


Table of contents

Concluding Remarks

- Deep learning showed unprecedented performance in different fields.
- It also showed competitive performance in wireless communication.
- 5G consists is envision to consist of unconventional system and technologies opening the opportunity to explore the power of Deep Learning.
- Deep learning can be used to
  - Replace several iterative schemes.
  - End to end systems.
  - Channel and non linearity models.
  - To provide schemes that may combine virtually nonhomogeneous side information.
- Deep Learning for propagation channel can be used to improve channel prediction capabilities, and improve system design and speed.
- Our initial results in propagation channel show that it could capture complex channels with consistent accuracy.
- There are large number of open problems where deep learning can be employed.

Questions?

Thanks to: Daoud Burghal, Zhiyuan Jiang

Contact information

Andreas F. Molisch

Ph.D., FNAI, FAAAS, FIEEE, FIET, MAASc.
Solomon-Golomb – Andrew-and-Erna-Viterbi Chair Professor
Head, Wireless Devices and Systems (WiDeS) Group,
Viterbi School of Engineering
University of Southern California (USC)
Los Angeles, CA, USA

Email: molisch@usc.edu
Website: wides.usc.edu

References

[1] Wymeersch, Henk, et al. "A machine learning approach to ranging error mitigation for UWB localization." *IEEE Transactions on Communications* 60.6 (2012): 1719-1728.

[2] Riihijarvi, Janne, and Petri Mahonen. "Machine Learning for Performance Prediction in Mobile Cellular Networks." IEEE Computational Intelligence Magazine 13.1 (2018): 51-60.

[3] Wang, Yong, Margaret Martonosi, and Li-Shiuan Peh. "Predicting link quality using supervised learning in wireless sensor networks." *ACM SIGMOBILE Mobile Computing and Communications Review* 11.3 (2007): 71-83.

[4] Wang, Jun-Bo, et al. "A Machine Learning Framework for Resource Allocation Assisted by Cloud Computing." *arXiv preprint arXiv:1712.05929* (2017).

[5] Wang, Yuyang, Murali Narasimha, and Robert W. Heath Jr. "MmWave Beam Prediction with Situational Awareness: A Machine Learning Approach." *arXiv preprint arXiv:1805.08912*(2018).

[6] G. Cybenko, “Approximation by superpositions of a sigmoidal function,” Math. of Control, Signals, and Systems, vol. 2, no. 4, pp. 303–314, 1989.

[7] Ibnkahla, Mohamed. "Applications of neural networks to digital communications–a survey." *Signal processing* 80.7 (2000): 1185-1215.

[8] A. Fehske, J. Gaeddert, and J. H. Reed, “A new approach to signal classification using spectral correlation and neural networks,” in Proc. IEEE Int. Symp. New Frontiers in Dynamic Spectrum Access Networks (DYSPAN), 2005, pp. 144–150

[9] O'Shea, Timothy J., Tugba Erpek, and T. Charles Clancy. "Deep learning based MIMO communications." *arXiv preprint arXiv:1707.07980* (2017).

[10] Cerri, G., et al. "Feed forward neural networks for path loss prediction in urban environment." *IEEE Transactions on Antennas and Propagation* 52.11 (2004): 3137-3139.

[11] Neskovic, Aleksandar, Natasa Neskovic, and Dore Paunovic. "Indoor electric field level prediction model based on the artificial neural networks." *IEEE Communications Letters* 4.6 (2000): 190-192.

[12] Samuel, Neev, Tzvi Diskin, and Ami Wiesel. "Learning to detect." *arXiv preprint arXiv:1805.07631* (2018).

[13] Farsad, Nariman, and Andrea Goldsmith. "Detection algorithms for communication systems using deep learning." *arXiv preprint arXiv:1705.08044* (2017).

[14] Farsad, Nariman, and Andrea Goldsmith. "Neural network detection of data sequences in communication systems." *arXiv preprint arXiv:1802.02046* (2018).

References

[15] Samuel, Neev, Tzvi Diskin, and Ami Wiesel. "Deep MIMO detection." *Signal Processing Advances in Wireless Communications (SPAWC), 2017 IEEE 18th International Workshop on*. IEEE, 2017.

[16] Burghal, Daoud, and Andreas F. Molisch. "Rate and Outage Probability In Dual Band Systems With Prediction-Based Band Switching." *IEEE Wireless Communications Letters* (2018).

[17] Zhao, Ming, and Wenye Wang. "WSN03-4: A Novel Semi-Markov Smooth Mobility Model for Mobile Ad Hoc Networks." *Global Telecommunications Conference, 2006. GLOBECOM'06. IEEE*. IEEE, 2006.

[18] E. Nachmani, Y. Be'ery, and D. Burshtein, "Learning to decode linear codes using deep learning," in Proc. Communication, Control, and Computing (Allerton), 2016, pp. 341–346

[19] Navabi, Shiva, et al. "Predicting Wireless Channel Features using Neural Networks." *arXiv preprint arXiv:1802.00107*(2018).

[20] He, Hengtao, et al. "Deep Learning-based Channel Estimation for Beamspace mmWave Massive MIMO Systems." *arXiv preprint arXiv:1802.01290* (2018).

[21] Ye, Hao, Geoffrey Ye Li, and Biing-Hwang Juang. "Power of deep learning for channel estimation and signal detection in OFDM systems." *IEEE Wireless Communications Letters* 7.1 (2018): 114-117.

[22] Burghal, Daoud, Rui Wang, and Andreas F. Molisch."Band Assignment in Dual Band Systems: Deep Learning-based Approach" MILCOM 2018.

[23] Burghal, Daoud, and Andreas F. Molisch."Deep learning for Robust Communication in Dual Band Systems" to be submitted 2018.

[24] Popescu, Ileana, et al. "Comparison of ANN Based Models for Path Loss Prediction in Indoor Environment." Vehicular Technology Conference, 2006. VTC-2006 Fall. 2006 IEEE 64th. IEEE, 2006.

[25] Östlin, Erik, Hans-Jürgen Zepernick, and Hajime Suzuki. "Macrocell path-loss prediction using artificial neural networks." IEEE Transactions on Vehicular Technology 59.6 (2010): 2735-2747.

[26] Neskovic, Aleksandar, Natasa Neskovic, and Djordje Paunovic. "Macrocell electric field strength prediction model based upon artificial neural networks." IEEE Journal on selected areas in Communications 20.6 (2002): 1170-1177.

[27] Cavalcanti, Bruno J., et al. "A Hybrid Path Loss Prediction Model based on Artificial Neural Networks using Empirical Models for LTE And LTE-A at 800 MHz and 2600 MHz." Journal of Microwaves, Optoelectronics and Electromagnetic Applications 16.3 (2017): 708-722.

[28] Timoteo, Robson DA, Daniel C. Cunha, and George DC Cavalcanti. "A proposal for path loss prediction in urban environments using support vector regression." Proc. Advanced Int. Conf. Telecommun. 2014.

References

[29] Zhang, Yan, et al. "Air-to-Air Path Loss Prediction Based on Machine Learning Methods in Urban Environments." *Wireless Communications and Mobile Computing* 2018 (2018).

[30] Jiang, Z., and Z. He, S. Chen, A. F. Molisch, S. Zhou, and Z. Niu, “Inferring Remote Channel State Information: Cramér-Rao Lower Bound and Deep Learning Implementation”, *Globecom 2018*, accepted.

Application in Wireless – End-to-End Systems [9]

- Use multi layer NN to perform joint modulation and coding.
- The system is an autoencoder that find a “compressed” efficient representation of the data at the transmitter side.
- Learns to (...)e “channel”.
- Was able t(...)ding with competitiv(...)