Deep Networks for Equalization in Communications

by

Laura Brink

A thesis submitted in partial satisfaction of the
requirements for the degree of
Master's of Science

in

Electrical Engineering and Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Anant Sahai, Chair
Professor John Wawrzynek

Fall 2018

The thesis of Laura Brink, titled Deep Networks for Equalization in Communications, is approved:

Chair	______________________________	Date	____________
	______________________________	Date	____________
	______________________________	Date	____________

University of California, Berkeley

Deep Networks for Equalization in Communications

Abstract

Deep Networks for Equalization in Communications

by

Laura Brink

Master's of Science in Electrical Engineering and Computer Science

University of California, Berkeley

Professor Anant Sahai, Chair

We apply the techniques from meta-learning to the communications domain. Specifically, we explore how equalization techniques can learn how to handle new environments without training on them.

Contents

List of Figures

Chapter 1

Introduction

With the rise of internet of things, more and more users will need to access the wireless spectrum. As the spectrum becomes saturated with users, not only must our systems adapt to changing environments but also be able to coexist and thrive in the presence of unknown neighbors. We must design a robust communications system that will give our devices freedom.

From an academic perspective, the rise of machine learning tools and processing has allowed us to tackle problems we have not yet been able to, like in image processing. However, we still do not fully understand the reach or the limitations of this technology. In order to study the limitations of machine learning, we must apply them to spaces that we have studied extensively, like communications, and compare them to the well-known baselines.

Most communications systems have three main processes; equalization, demodulation, and error-correction. While we will need to design robust forms of all of these processes, we will focus on equalization for the remainder of this paper.

1.1 Motivation

Why are we doing this?

What is meta-learning?

[5]

[9]

1.2 Background

Inter-symbol Interference and Equalization

Inter-symbol interference occurs when we are transmitting over a channel that has some echos. These echos cause the receiver to hear a garbled signal instead of the original signal from the transmitter. This is called inter-symbol interference because the receiver is hearing a combination of symbols across time.

Let $\vec{x} = [x_0, x_1, \dots x_n]$ be the set of n complex symbols that the transmitter sends over the channel that connects the transmitter to the receiver. Each channel will have different characteristics. Some channels may have echos, others may have delays, often channels will have both. When a channel has echos, this is called a multipath channel because there are multiple paths to reach the receiver. Each path is called a tap. We can characterize a channel by characterizing the taps.

What can cause these echos? How prevelant are they?

Channel taps: Let $\vec{a} = [a_0, a_1, \dots a_l]$ be the set of characteristic for a multipath channel that has l taps. When a sequence of symbols like $\vec{x}$ is transmitted over this channel, the channel taps are convolved over the sequence. Additionally, there is noise in the system denoted by η_i.

$$\tilde{x}_m = \sum_{i=0}^{l} a_i x_{m-i} + \eta_i$$

The receiver will hear a signal that is corrupted by inter-symbol interfence and noise; $\tilde{\vec{x}} = [\tilde{x}_0, \tilde{x_1}, \dots \tilde{x}_{n+l}]$.

Receivers must be able to handle garbled signals in order to transmit data in the real world. The process of removing the inter-symbol interference is called equalization. The goal of equalization is to take in a garbled signal and output a signal with minimal inter-symbol interference.

INSERT IMAGE OF 2 TAP CHANNEL EFFECTS ON QPSK

Engineers have built processes to remove inter-symbol interference. First, let's go into the case when the channel characteristics are known.

Equalization for a known channel

If you know the channel characteristics, $\vec{a}$, perfectly, then there are a few different methods that can be used.

Zero-forcing

While it's important to consider how well a receiver can equalize with a known channel, this is rarely the case. Usually, we do not know the channel characteristics.

Equalization for an unknown channel

When the receiver does not know the channel characteristics, the process of equalization essentially has two jobs; first, identify the channel, second, remove the inter-symbol interference. If the receiver did not identify the channel first, there would be no way to remove the affects of it on the received signal.

In order to do channel estimation, most systems require that packets begin with a known sequence called a preamble.

Channel estimation: least squares
Minimum mean squared error equalizer.

How do real systems handle equalization?

OFDM does not have this problem!

Carrier Frequency Offset and Correction

Now, if we were to implement our minimum mean squared error algorithm on a physical receiver, we would find some problems with our equalization process. Our equalizer will equalize the first symbols very well. However, as we equalize end parts of our sequence, we will encounter a physical phenomonen called carrier frequency offset, CFO.

Carrier frequency offset occurs when ???

When there is a significant CFO present, our symbols will gradually start rotating. CFO will effect our received symbols like

$$\tilde{x}_m = x_m e^{mj\omega}$$

The effect will look like something like this
INSERT IMAGE OF CFO OCCURRING on QPSK

How do real systems handle CFO correction?

There are a few ways to handle CFO, some are more elegant than others.

The first solution is to try to remove the problem. Since CFO is dependent on the length of a packet, one solution is to make packets so short that

A more elegant solution is using phase-lock loops (costas loops).

What must a modern day receiver handle? What does it look like when we have both CFO and ISI?

$$\tilde{x}_m = (\sum_{i=0}^{l} a_i x_{m-i}) e^{mj\omega} + \eta_i$$

INSERT IMAGE OF AFFECTS OF BOTH CFO AND EQUALIZATION
INSERT IMAGE OF SAME THING WITH NOISE!

1.3 Related Works

Recently, many researchers have become more interested in applying deep learning techniques to communications systems. We refer the reader to [1][2][17][6] for surveys of these kind of works. Most of the works that we have found typically only deal with additive white gaussian noise (AWGN) channels. Only dealing with AWGN channels, essentially removes the problem of equalization.

In [3], Dörner et. al. implemented an end to end transmitter and reciever with neural networks that allowed gradients to flow all the way back from the receiver during training. However, they restricted themselves to only sending a certain set of messages, only seeing AWGN channels, and they did not address CFO.

Other groups have been focusing on decoding with recurrent neural networks (RNNs) but also only deal with AWGN channels [7][8]. Others have been working with generative adversarial networks (GANs) to train an end-to-end communication system [18]. However, they also only consider AWGN channels or Rayleigh fading channels.

For those that do consider more complex channels, most re-train their models for each new channel seen. Ye et. al. consider OFDM systems where their feedforward neural networks estimate the channel state information then train offline for that specific channel [19]. [16] considers nonlinear channels but re-trains their network for each new channel. Note, this work does not go into detail about the architecture used or how the networks are trained.

Optics and molecules? [4]

Timothy O'Shea's group has been doing some excellent work in this area. In [15], they use convolutional neural networks (CNNs) to estimate, but not correct, carrier frequency offset and timing offset. However, they only consider AWGN channels and Rayleigh fading channels. In [13], they explore how unsupervised learning can train autoencoders for multiple antenna communications. They do re-train for each new channel and use a Rayleigh fading channel model.

[11] [12] [14] [10]

Chapter 2

Deep Networks for Equalization

2.1 Replicate Results

2.2 Channel Estimation

- compare least squares and how KNN did with pure deep nn based architecture
- NN did better than least squares
- hyperparam search over general 1-layer to 4-layer dense layers, and number of nodes and activations
- plots: how error changed with respect to data points, as number of data points increased, NN outperformed LS
- preamble: 100
- want: QPSK, plot of preamble length
- want? how to visualize that NN does better than LS

2.3 Channel Equalization

- compare to MMSE
- re run with the new RNN architecture
- backprop length of 3. crude search from 1-10. 2 tap channel
- added channel preprocessing: didn't seem to make too much of a difference
- plot log/ log scale to find converging in error

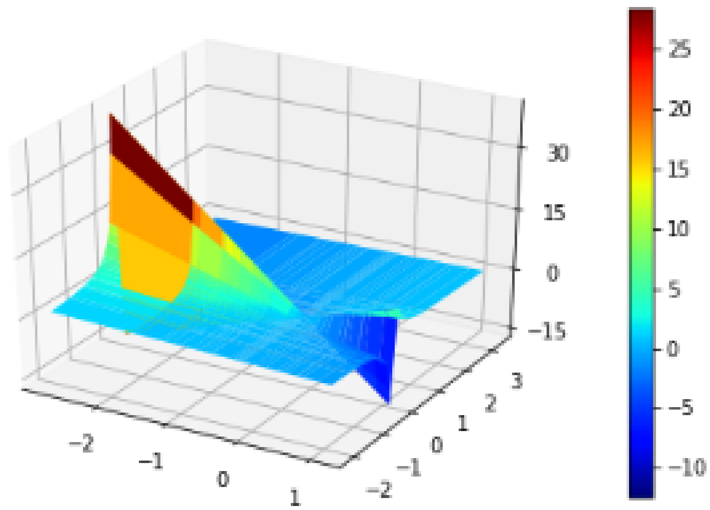

Figure 2.1: Division function; $z = \frac{x}{y}$.

Learning an inverse

- can a NN learn to do division?
- given a one tap channel and data sequence, output is equalized data sequence
- with and without log feature scaling
- without log errors = 10^-6
- MMSE gets error = 10^-32
- with log errors =
- straight inversion, without log feature scaling 10^-7 - with dense layers
- inversion with log feature scaling with error of 10^-14
- plots: inversion, but as a function of beta for both non-log and log

Learning to multiply two inputs

NN given x, y - output x*y

2.4 Channel Est + Equal

re run that

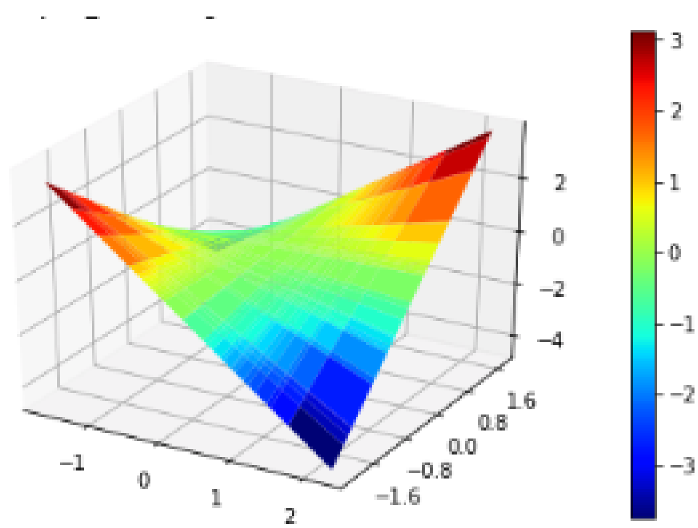

Figure 2.2: Multiplication function; $z = xy$.

Chapter 3

Deep Networks for CFO

3.1 Recurrent Neural Network Follows a Circle

for a constant rate for a given rate

3.2 Deep Network Carrier Frequency Offset Estimation

- complex gradients problems
- act like the real and imaginary parts are separate
- plots: one tap channel plots, without equalization problems
- plots: two tap channel plots, with equalization problems

3.3 Deep Network Carrier Frequency Offset Correction

Program a Costas loop for comparison

Chapter 4

Conclusion

conclucsion

Seria Electronică Şi Telecommuncaţii, Transactions on Electronics and Communications

Bibliography

[1] Corina Botoca, Georgeta Budura, and Miclau Nicolae. "Nonlinear Channel Equalization Using Complex Neural Networks". In: *Seria Electronică Şi Telecommuncaţii, Transactions on Electronics and Communications* (2004), pp. 227–231.

[2] Theo Diamandis. "Survey on Deep Learning Techniques for Wireless Communications". In: *Departement of Electrical Engineering, Stanford University* (2017).

[3] Sebastian Dörner et al. "Deep Learning-Based Communication Over the Air". In: *arXiv preprint arxiv:1707.03384* (2017).

[4] Nariman Farsad and Andrea Goldsmith. "Neural Network Detection of Data Sequences in Communication Systems". In: *IEEE Transactions on Signal Processing* PP (2018).

[5] Chelsea Finn, Pieter Abbeel, and Sergey Levine. "Model-agnostic meta-learning for fast adaptation of deep networks". In: *International Conference on Machine Learning* (2017).

[6] Hengtao He et al. "Model-Driven Deep Learning for Physical Layer Communications". In: *arXiv preprint arxiv:1809.06059* (2018).

[7] Hyeji Kim et al. "Communication Algorithms via Deep Learning". In: *International Conference on Learning Representations* (2018).

[8] Hyeji Kim et al. "Deepcode: Feedback Codes via Deep Learning". In: 2018.

[9] Brenden M. Lake, Ruslan Salakhutdinov, and Joshua B. Tenenbaum. "Human-level concept learning through probabilistic program induction". In: *Science*. Vol. 350. 6266. 2015, pp. 1332–1338.

[10] Timothy J. O'Shea, Kiran Karra, and T. Charles Clancy. "Learning to Communicate: Channel Auto-encoders, Domain Specific Regularizers, and Attention". In: *IEEE International Symposium on Signal Processing and Information Technology*. 2016, pp. 223–228.

[11] Timothy J. O'Shea, Tamoghna Roy, and Nathan West. "Approximating the Void: Learning Stochastic Channel Models from Observation with Variational Generative Adversarial Networks". In: *arXiv preprint arxiv:1805.06350* (2018).

[12] Timothy J. O'Shea et al. "Radio Transformer Networks: Attention Models for Learning to Synchronize in Wireless Systems". In: *arXiv preprint arxiv:1605.00716* (2016).

[13] Timothy O'Shea, Tugba Erpek, and T. Charles Clancy. "Deep Learning-Based MIMO Communications". In: *arXiv preprint arXiv:1707.07980* (2017).

[14] Timothy O'Shea and Jakob Hoydis. "An Introduction to Deep Learning for the Physical Layer". In: *IEEE Transactions on Cognitive Communications and Networking*. Vol. 3. 4. 2017, pp. 563–575.

[15] Timothy O'Shea, Kiran Karra, and T. Charles Clancy. "Learning Approximate Neural Estimators for Wireless Channel State Information". In: *arXiv preprint arXiv:1707.06260* (2017).

[16] K. T. Raghavendra and Amiya K. Tripathy. "An Efficient Channel Equalizer Using Artificial Neural Networks". In: *Neural Network World* 16 (2006), pp. 357–368.

[17] Tianqi Wang et al. "Deep Learning for Wireless Physical Layer: Opportunities and Challenges". In: *China Communications* 14.11 (2017), pp. 92–111.

[18] Hao Ye, Geoffrey Ye Li, and Biing-Hwang Juang. "Channel Agnostic End-to-End Learning based Communication Systems with Conditional GAN". In: *IEEE Global Communications Conference*. 2018.

[19] Hao Ye, Geoffrey Ye Li, and Biing-Hwang Juang. "Power of Deep Learning for Channel Estimation and Signal Detection in OFDM Systems". In: *IEEE Wireless Communications Letters*. 2018, pp. 114–117.