Parameter Estimation in an Optical Fibre System Using Deep Learning

Yuvi Dhelawat (1801EE62)
Supervisor – Dr. Sumanta Gupta
Department of Electrical Engineering, Indian Institute of Technology, Patna

1. Objective

To train a deep learning model for the estimation of Channel length and the type of modulation of signal to aid in extraction of information more faithfully in an Optical fibre system.

2. Need in Today's world

In modern days, for a dynamic reconfigurable optical fibre network, the interconnections between two nodes vary based on the demand. At the transmitter end, one cannot predict when there will be a connection between two nodes. For two nodes A and B to be connected at different times, the length of the path may vary. Also, modulation format and the data rate can be switched based on the type of data one wants to send. This approach proves to be efficient but for information to be extracted from the received data, more insight on the path, modulation scheme, bit rate, etc. is needed for a better extraction.

Our model, if placed at the receiver end, can extract the parameters of the system such as the link length and the modulation scheme used which can be further used to employ signal processing techniques for a more faithful extraction of information.

3. DEEP LEARNING

Ever since the days of inception, the usage of Deep Learning techniques has always been expanding to broader domains. The idea behind these seemingly-invincible techniques is quite simple to understand, it is to replicate the way our brain functions. Our brain is a cluster of neurons, each transmitting data to its neighbours using small electrical discharges. The raw data is processed as it passes through the complex network of neurons, and at each stage, some information is extracted. These all combined help at creating a decision. The mathematical realization replaces these neurons with matrices, while the complex network remains intact. The inner matrices of each of the neurons are updated by back-propagating the error measured at the output layer. Thus, the model always tries to converge at the least possible value of error.

Now, given the diverse usage deep learning has been demonstrating, its usage in Optical Fibre systems has been explored vastly in recent decades. Earlier approximation techniques applied had to breakdown the entire system in a block-wise manner (receiver, channel, transmitter). Thus, extensive mathematical knowledge of the system was needed to design the model. However, the capacity of deep learning models to approximate any nonlinear function has helped researchers in creating a more robust estimation system without actually approximating the system in a block-wise manner, rather doing the estimation in a single end to end process. This enables a communication system to be flexibly adaptive for information transmission over any type of channel without requiring prior mathematical modelling and analysis. The viability of such models has been extensively explored in communication systems.

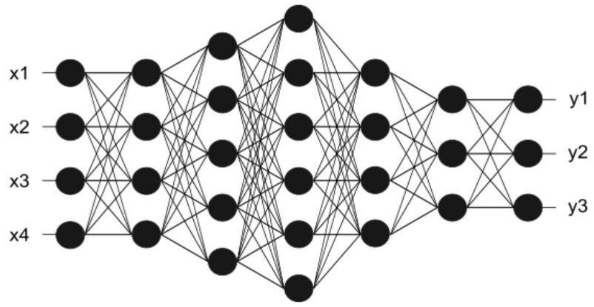

Fig. 1 Basic Structure of a Neural Network

4. Why CNN

In our approach we have preferred using a Convolutional Neural Network for estimation of modulation technique as well as the channel length for these major reasons:

1. CNNs have been complimented for their expertise in image classification. Now, in our case, even though we trained our model for the numerical values of the received signals, it could be equivalently described as using the constellation diagrams for training the model, i.e., using an image to classify the modulation technique and regress the length of the channel.
2. Given that two convolutional layers are never fully connected in a CNN, they prove to be computationally less expensive as compared to their Dense counterparts.
3. Given a dataset, CNNs inherently perform feature selection, convolving the data with various filter matrices at each layer.

However, there are certain disadvantages of using a CNN as well:

1. Given the partially connected layers, CNNs are less effective when it comes to training on data points as they inherently tend to extract important features. In

a dataset where each value carries equal importance, CNN fails to outperform DNN.
2. Added layers for improving functionality such as MaxPooling and Dropouts tend to make CNN slower at training.
3. Larger dataset is needed usually for training a CNN.

Thus sticking to a middle ground and using a model with precisely the right amount of convolution and dense layers helped at achieving smaller losses at a fast training speed.

5. Communication System

A communication system transmits information from one place to another, whether separated by a few kilometers or by transoceanic distances. Information is often carried by an electromagnetic carrier wave whose frequency can vary from a few megahertz to several hundred terahertz. Optical communication systems use high carrier frequencies (∼100 THz) in the visible or near-infrared region of the electromagnetic spectrum.

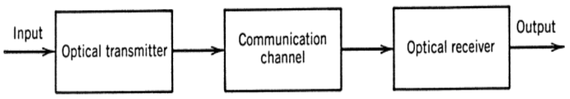

Fig. 2. General Block Diagram for an optical communication system

6. Optical Fibres as a Communication Channel

The role of a communication channel is to transport the optical signal from transmitter to receiver without distorting it. Most lightwave systems use optical fibres as the communication channel because silica fibres can transmit light with losses as small as 0.2 dB/km. Even then, optical power reduces to only 1% after 100 km. For this reason, fibre losses remain an important design issue. Another important design issue is fibre dispersion, which leads to broadening of individual optical pulses with propagation. If optical pulses spread significantly outside their allocated bit slot, the transmitted signal is severely degraded. Eventually, it becomes impossible to recover the original signal with high accuracy. The problem is most severe in the case of multimode fibres, since pulses spread rapidly (typically at a rate of ∼10 ns/km) because of different speeds associated with different fibre modes. It is for this reason that most optical communication systems use single-mode fibres.

7. Impairments in an Optical Fibre System

Major impairments in an Optical Fibre system can be broadly classified into two categories:

7.1. Linear Impairments

1. Attenuation –
It causes the signal strength to decay as it propagates through the fibre, caused by intrinsic factors such as scattering and absorption and extrinsic factors such as stress from the manufacturing process, environmental and physical bending.
2. Chromatic Dispersion –
Due to this, different wavelengths travel at different speeds, causing the pulse to spread out. If left unmanaged, the pulses spread out so much that they cause inter-symbol interference, making it hard to separate and extract information.

7.2. Non-linear Impairments

When an optical fibre system is operated at higher bit rate, greater than 10GBPS and at higher power, non-linearity become dominant in the system. The dependence of the refractive index on square of field amplitude causes the non-linearity to arise. At sufficiently high optical intensities, non-linear refraction occurs (Kerr effect), which is basically the variation of the refractive index with light intensity. In a single mode fibre, even a single light wave can be affected by this nonlinearity since its phase is modulated by optical intensity fluctuations in the same wave.

8. Optical Fibre System Parameters

8.1. OSNR

Optical signal-to-noise ratio is used to quantify the degree of optical noise interference on optical signals. It is the ratio of service signal power to noise power within a valid bandwidth. It is a key parameter to estimate performance of Optical Networks.

8.2. Pulse Peak Power

The power of an optical pulse varies with time and its maximum is called the peak power.

8.3. Modulation Frequency

It is the frequency used for the modulation of signals during communication.

8.4. Channel Length

It is the length of the optical fibre used for the transmission of signal from transmitter to receiver.

8.5. Modulation of signal

Modulation is the process of varying one or more properties of a periodic waveform, called the carrier signal, with a separate signal called the modulation signal that typically contains information to be transmitted.

9. MODULATION FORMATS USED

9.1. OOK (On Off Keying) – It denotes the simplest form of ASK(Amplitude - Shift Keying) in which the signal power is simply changed between two levels , one of which is set to zero , to reflect the on-off nature of the resulting optical signal.

9.2. BPSK (Binary Phase Shift Keying) – It is a two phase modulation scheme, where the 0's and 1's in a binary message are represented by two different phase states that have a phase difference of 180 degrees in the carrier signal.

9.3. QPSK (Quadrature Phase Shift Keying) – It is a form of Phase Shift Keying in which two bits are modulated at once, selecting one of four possible carrier phase shifts (0, 90, 180, or 270 degrees). QPSK allows the signal to carry twice as much information as ordinary PSK using the same bandwidth.

10. CONSTELLATION DIAGRAMS FOR VARYING PARAMETERS

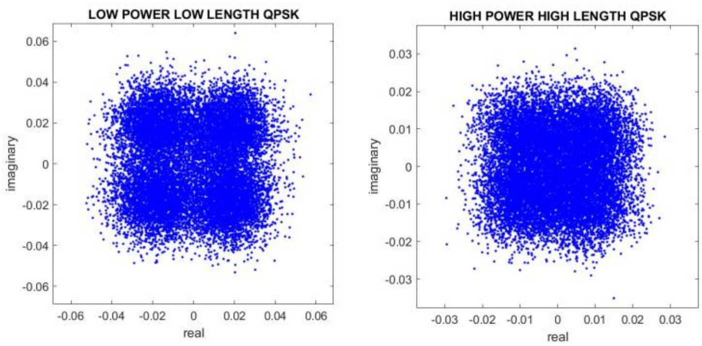

Fig. 3 a) Low Power Low Length QPSK b) High Power High Length QPSK

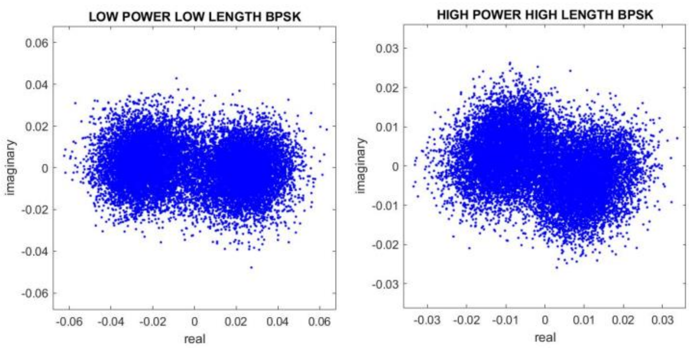

Fig. 4. a) Low Power Low Length BPSK b) High Power High Length BPSK

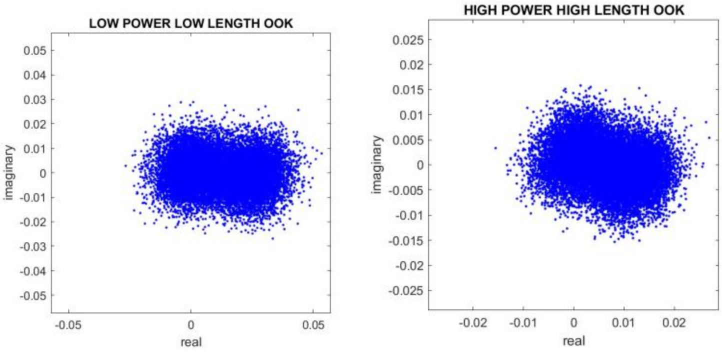

Fig. 5. a) Low Power Low Length OOK b) High Power High Length OOK

Clearly, the clusters of data-points tend to lose their identity as we go on increasing the pulse peak power and the length of the channel. The increase in pulse peak power tends to increase the nonlinearity in the system, which in turn causes distortion in phase and this corruption can be clearly seen in the constellation diagram. The length on the other end helps attenuation and chromatic distortion to dominate.

11. MODEL TRAINING AND TESTING

The models were trained by taking the entire system as a black box. The bits received on the receiver end were labelled with corresponding parameters, i.e., channel length and type of modulation. These were the values used for creating the dataset:

- Signal to Noise Ratio - 15 (in dB)
- Modulation Frequency - 10GBPS
- Pulse Peak Power - 1 (in dBm)
- SMF Length - {5, 10, 15, 20… 75, 80} (in kms)
- Type of modulation - OOK, BPSK, QPSK
- Number of bits per signal - 1000
- Number of samples taken per bit - 16

A larger number of bits per signal was preferred so as to have a clear clustering of points in the constellation diagram, for better estimations using CNN. This data was split into training, validation and testing data in 3:1:1, where 21600 signals were used in training and 7200 signals each were used for testing and validation. A larger training dataset helped at ensuring that that model was trained for all the cases effectively, while a large validation dataset helped at ensuring that there was no over-fitting.

The model architecture for classification model had 32, 32, 64 neurons in the convolution layer followed by a flattening layer and then dense layers having 256, 128, 64, 3 neurons. Relu activation was preferred for all the layers except the last one, which had softmax activation. Maxpooling, BatchNormalization and dropout layers were added after each convolution layer to assist at better training. The model was compiled with mean absolute error loss and adam optimizer.

The regression model had 32, 128, and 256 neurons in the convolution layer followed by a flattening layer and then dense layers having 256, 128, 1 neuron-layers. Relu activation was preferred for all the layers. Maxpooling and dropout layers were added after each convolution layer to assist at better training. The model was compiled with mean absolute error loss and adam optimizer. A callback was used to reduce the learning rate whenever the model hit a plateau.

These model, were then tested on 75 different combinations of parameters:

- Fm - {10, 20, 40} (in GBPS)
- OSNR - {10, 15, 20, 25, 30} (in dB)
- Pulse Peak Power - {0, 1, 3, 5, 8} (in dBm)
- Length of the Channel - {5, 10, 15… 75, 80} (in Kms)
- Type of Modulation - OOK, BPSK, QPSK

Each dataset had 384 samples and was used purely for testing both models. The results were:

	OSNR	Fm	Pulse Peak Power	Classification Accuracy	regression loss
0	10	10	0	0.992126	3.142066
1	10	10	1	0.992126	4.558811
2	10	10	3	1.000000	12.226591
3	10	10	5	0.994751	19.142027
4	10	10	8	0.989501	27.312475
5	10	20	0	0.992126	3.468806
6	10	20	1	0.997375	4.871821
7	10	20	3	0.986877	12.563801
8	10	20	5	0.992126	19.556753
9	10	20	8	0.992126	27.618806
10	10	40	0	0.939633	3.828895
11	10	40	1	0.950131	3.410819
12	10	40	3	0.921260	10.768244
13	10	40	5	0.855643	17.785129
14	10	40	8	0.847769	26.134871
15	15	10	0	1.000000	6.917360
16	15	10	1	1.000000	3.307008
17	15	10	3	1.000000	7.082564
18	15	10	5	1.000000	15.101402
19	15	10	8	1.000000	24.404130
20	15	20	0	1.000000	6.327225
21	15	20	1	1.000000	3.027951
22	15	20	3	1.000000	7.700180
23	15	20	5	1.000000	15.583832
24	15	20	8	1.000000	24.830724
25	15	40	0	1.000000	6.657905
26	15	40	1	1.000000	3.383113
27	15	40	3	1.000000	7.329561
28	15	40	5	0.984252	14.838875
29	15	40	8	0.984252	24.410839
30	20	10	0	1.000000	9.492815
31	20	10	1	1.000000	5.335812
32	20	10	3	1.000000	4.548225
33	20	10	5	1.000000	12.776912
34	20	10	8	1.000000	23.062451
35	20	20	0	1.000000	8.726571
36	20	20	1	1.000000	5.028674
37	20	20	3	1.000000	5.297308
38	20	20	5	1.000000	13.634361
39	20	20	8	1.000000	23.668128
40	20	40	0	1.000000	8.135971
41	20	40	1	1.000000	4.140118
42	20	40	3	1.000000	5.895406
43	20	40	5	1.000000	13.824677
44	20	40	8	1.000000	23.710712
45	25	10	0	1.000000	10.482712
46	25	10	1	1.000000	6.571389
47	25	10	3	1.000000	3.923262
48	25	10	5	1.000000	11.926109
49	25	10	8	1.000000	22.504972
50	25	20	0	1.000000	9.756888
51	25	20	1	1.000000	5.878666
52	25	20	3	1.000000	4.641349
53	25	20	5	1.000000	12.848618
54	25	20	8	1.000000	23.248347
55	25	40	0	1.000000	8.643042
56	25	40	1	1.000000	4.537920
57	25	40	3	1.000000	5.350817
58	25	40	5	1.000000	13.500459
59	25	40	8	1.000000	23.356254
60	30	10	0	1.000000	10.846488
61	30	10	1	1.000000	6.933140
62	30	10	3	1.000000	3.897178
63	30	10	5	1.000000	11.522279
64	30	10	8	1.000000	22.311249
65	30	20	0	1.000000	10.057626
66	30	20	1	1.000000	6.186714
67	30	20	3	1.000000	4.485686
68	30	20	5	1.000000	12.509805
69	30	20	8	1.000000	22.945043
70	30	40	0	1.000000	8.767692
71	30	40	1	1.000000	4.784634
72	30	40	3	1.000000	5.230405
73	30	40	5	1.000000	13.272576
74	30	40	8	1.000000	23.332877

Table 1. Testing Results

While the classification model showed good results on each of the cases, the regression model had difficulties with the cases it wasn't trained on.

So, for the estimation of channel length, the testing dataset was used for training the model, having it learn from received bits labelled with parameters such as Fm, OSNR, Pulse Peak Power, SMF Length and the type of modulation. The model had a single convolution layer at the top with 64 neurons followed by a flattening layer and then a series of dense layers having 1024, 512, 256, 200, 128, 64, 32, 20, 10, 1 neuron-layers, each with relu activation. This model showed better results with a testing loss of just 1.94. To understand the efficiency of this model, let's take a look at a few estimated values:

	OSNR	Fm	Pulse Peak Power	Expected Length	Estimated Length
0	[10.0]	[1.0]	[0.0]	[80]	79.599998
1	[20.0]	[4.0]	[8.0]	[70]	66.500000
2	[20.0]	[2.0]	[3.0]	[80]	77.699997
3	[30.0]	[2.0]	[1.0]	[30]	30.000000
4	[30.0]	[1.0]	[1.0]	[70]	70.099998
5	[15.0]	[2.0]	[5.0]	[30]	30.000000
6	[15.0]	[1.0]	[0.0]	[70]	71.500000
7	[20.0]	[4.0]	[5.0]	[60]	54.400002
8	[10.0]	[4.0]	[0.0]	[80]	80.099998
9	[25.0]	[4.0]	[3.0]	[10]	10.000000
10	[30.0]	[2.0]	[1.0]	[50]	50.200001
11	[20.0]	[4.0]	[8.0]	[10]	9.800000
12	[20.0]	[2.0]	[5.0]	[20]	19.500000
13	[25.0]	[1.0]	[0.0]	[40]	40.299999

Table 2. Few estimated Values for improved Regression Model

While most of the estimated values stand close to the real ones, a large error is observed in case#1, which can be due to larger Pulse Peak Power which increases the non-linearity of the system. Nonetheless, the model worked efficiently for most of the cases, as can be seen in the image above.

12. CONCLUSION

1. The classification model was able to predict the type of modulation format of signal (here we used three formats i.e. OOK, BPSK and QPSK) with good accuracy, when trained only with the received signal bits as parameters.

2. The first regression model didn't perform well when trained only with the received signal bits as parameters and the probable reason may be the increase in non-linearity in fibre as the pulse peak power increases, due to which phase of the signal gets distorted and the signal is corrupted.

3. But, the second regression model was more robust and gave marginal loss, however, it needed additional information about the original signal, i.e., OSNR, Fm, and Pulse Peak Power.