Machine Learning Based Optimal Modulation Format Prediction for Physical Layer Network Planning

Danish Rafique
ADVA Optical Networking SE, Fraunhoferstr. 9a, 82152, Munich, Germany
e-mail: DRafique@ADVAoptical.com

ABSTRACT
Physical layer network design and planning process is a cumbersome one. It includes laying out all possible combinations of modulation formats, fiber types, forward error correction codes, channel spacing, etc., conducting exhaustive simulations and lab experiments to come up with carefully tuned engineering rules, and finally using these approximate models to propose transmission feasibility. Besides being cumbersome, there are two fundamental issues in conventional network planning approach, firstly it almost exclusively offers conservative design, leading to resource underutilization, and secondly it's not scalable – neither from planning viewpoint nor computationally – to next-generation highly granular and flexible networks.

Machine learning, an artificial intelligence toolset, may be applied to solve aforementioned issues by allowing data-driven model development, and consequent transmission quality prediction. While network planning is an extensive topic, in this paper, we focus on neural network based modulation format classification, autonomously identifying best possible modulation format for a given link configuration.
Keywords: communication networks, machine learning, analytics, optimization, optical fiber communications.

1. INTRODUCTION

Optical networks are susceptible to various deteriorations over their lifetime, ranging from equipment failures to optical power variations, gain saturation behaviour, nonlinear channel response, etc. [1-3]. Consequently, the knowledge of these potential distortions needs to be accounted for in order to efficiently design and plan the network. Traditionally, this has been achieved using precise analytical models, augmented by engineering rules devised from exhaustive simulation and laboratory experiments, and fine-tuned to the network at hand. These approximated analytical approaches led to resource underutilization (conservative planning due to intrinsic inaccuracies), high computational loads, and no real-time visibility on network state. Furthermore, as the complexity of modern networks has grown [4-6], the sheer number of optimization parameters has substantially increased. For instance, the choice of symbol rates, channel grid, modulation scheme, etc. makes it impossible to lay out precise models encompassing the available parametric space.

On the other hand, optical networks have evolved to be more programmable and centralized with the introduction of software-defined networking, separating the control and data plane [7-9]. SDN allows for unique network monitoring capabilities, enabling real-time insight into current network states. To this end, machine learning may be employed to enable data-driven network design and optimization, as opposed to expert features and decision criterions. These computational approach allows for real-time resource optimization, together with scalable models, updated as a function of network data [10-15]. Figure 1a shows a qualitative example of physical layer transmission capacity optimization. The conventional approach would identify and deploy a fixed capacity for the entire network lifetime using a look-up table (or a variant), whereas ML-driven solution enables dynamic capacity allocation based on real-time network conditions. Figure 1b shows workflow of such an approach, where network monitoring probes collect data, used to train a ML model, followed by both in-operation use of that model for optimum capacity prediction and also continuous model adaptation to real-time network data.

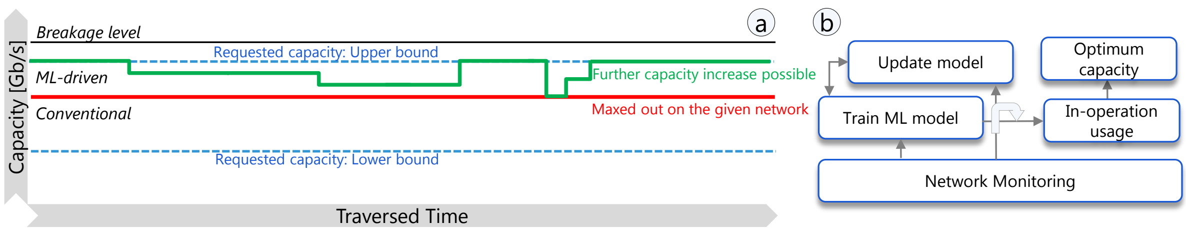


Figure 1: (a) Qualitative example of conventional versus machine learning based physical layer transmission capacity allocation;
(b) Typical workflow of machine learning model development and usage.

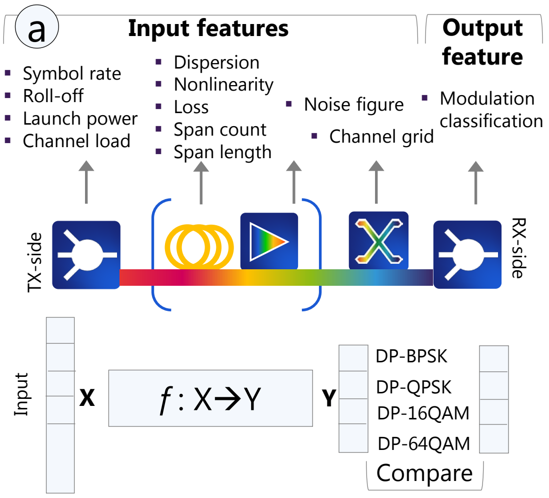
Recently published work has focused on predicting quality of transmission based on measured bit error rate (BER) and Q-factor [16-17], in this work we aim to rather predict optimum modulation format based on features such as symbol rate, channel load, number of spans, etc. In particular, we consider various multi-layer perceptron

(MLP) architectures, and report the performance in terms of classification accuracy and training time. Our results suggest that while simplistic single layered MLP may enable accuracy in excess of ~92%, a 2-layered MLP with 100×10 neurons enables accuracy of ~98% with a training time of ~20s.

2. Setup, Dataset and Classifier Details

Figure 1a shows the setup considered in this work. At the transmit-side, we considered typical features like channel symbol rate, pulse shaping filter roll-off, optimum launch power and channel load on a given point-to-point optical link. The optical transmission channel was modelled using the Gaussian noise (GN) approach described in [18]. The back-to-back optical signal-to-noise ratio (OSNR) penalties were considered to be 0.5dB, 1dB, 2dB and 3dB for dual polarization phase shift keying (DP-BPSK), DP-QPSK, dual polarization quadrature amplitude modulation (DP-16QAM), and DP-64QAM, respectively, accounting for optical and electrical component limitations. All of the considered features are enlisted in Fig. 1a. We used GN simulations to generate our well characterized dataset, where the training data consisted of 6×10^4 unique records, validation data consisted of 2×10^4 records, and test dataset comprised of 2×10^4 records. The validation data is used to choose the best model, avoiding overfitting to training data instances. The quantitative parameter value ranges are shown in Fig. 1b, representing extensive transmission scenarios. It is worth mentioning that the synthetic data generation model (GN) represents realistic transmission behaviour, as attested by several bodies of work, and the models created therein may be extended based on in-field datasets as they become available.

The artificial neural network (ANN) classifier was a feedforward model with supervised learning based on backpropagation algorithm, termed as multi-layer perceptron (MLP), and was trained on the 11 input features and 1 output feature listed in Fig. 1a. The output feature was a multi-level categorical variable representing the four modulation formats mentioned above. The goal of the classifier was to predict the true class, given a set of input features. The performance was evaluated using accuracy measure, i.e. the number of correct predictions given a class label. Finally, we considered four ANN configurations, ranging from a single layer with 5 neurons to 2 layers with 100 and 10 neurons each (see other values in Fig. 1b). The hidden layers used rectified linear activation functions, except the output layer which used soft-max activation functions. Note that we used rather shallow ANN architectures as one of our objectives is to establish performance versus MLP complexity trade-off for an optical network prediction task.



b

Training data size	60000
Valid/Test data size	20000/20000
Algorithm	MLP
Architectures	1:5 (MLP_A), 1:10, (MLP_B), 1:100, (MLP_C), 2:100×10 (MLP_D)
Batch size	100
Learning rate	0.01
Activation	ReLu

c

Symbol rate	Roll-off	Launch power	Channel load	Dispersion	Nonlinear index	Loss	Span count	Span length	Noise figure	Channel grid
30-90 Gbaud	0.01-0.05	f(symbol rate)	1-120	4-21 ps/nm/km	0.8-1.6 1/W.km	0.15-0.2 dB/km	1-50	40-120 km	4-6.5 dB	f(symbol rate)

Figure 2: (a) Optical link setup, and list of input and output features for MLP model development. Output feature consists of multi-level classes, representing DP-BPSK, DP-QPSK, DP-16QAM and DP-64QAM; (b) Details of training and test data, together with list of MLP parameters; (c) Quantitative ranges for input and output features used in this work.

3. RESULTS AND DISCUSSIONS

In order to evaluate performance of our classifiers, we evaluated their performance as a function of epochs, as shown in Fig. 3a. The curves represent optimization behaviour on validation data, as the test-data model was chosen based on convergence obtained on validation set. It can be seen that all of the MLP architectures follow a similar saturation behaviour after certain number of epochs. MLP_A achieves 93.6% accuracy after 6000 epochs, followed by MLP_B with 95.8% accuracy and 5000 epochs, MLP_C with 97.07% accuracy after 3000 epochs, and MLP_D with 97.36% accuracy only after 2000 epochs. Clearly, MLP_D enables a significant 5% accuracy gain with three time lower required number of epochs. Figure 3b shows corresponding elapsed time to obtain optimum model for the four classifiers. Interestingly, the least complex to the most complex configuration follows a linearly decreasing required model training time, ranging from ~75 s to ~20 s. This behaviour may be attributed to lower number of required epochs for the more complex models, ascertaining that the more complex architecture doesn't significantly contribute to the training time. Nonetheless, MLP_D achieves an impressive 97.36% classification accuracy, making it or even a more complex architecture a suitable choice for modulation classification problem in optical networks.

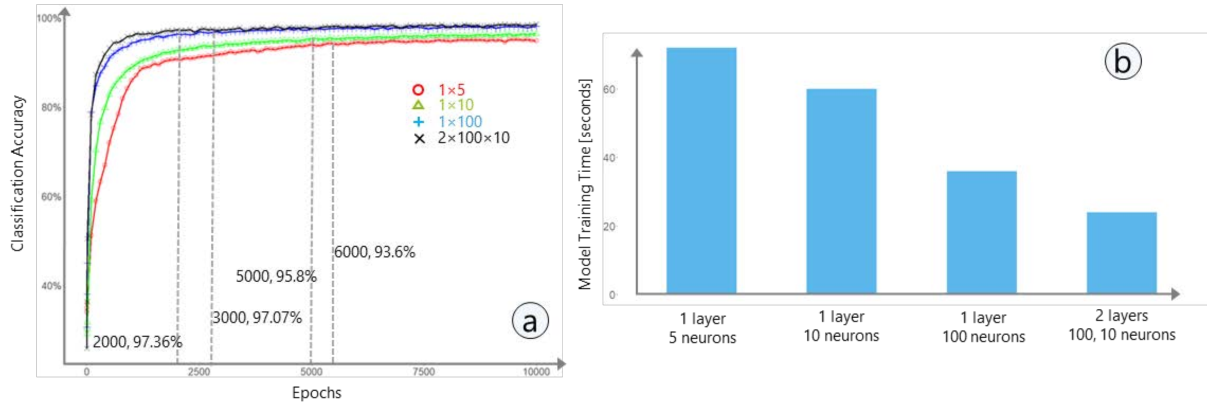


Figure 3: (a) Classification accuracy as a function of number of epochs for MLP_A (circles), MLP_B (triangles), MLP_C (plus), and MLP_D (cross); (b) Model training time to reach convergence in Fig. 3a for the four MLP architectures considered.

Figure 4 provides further details on results, evaluated on test data, showing detailed classification results as a function of modulation categories for the four MLP structures. Fig. 4a depicts that the lowest classification accuracy for MLP_A is largely attributable to misclassification of DP-BPSK as DP-QPSK, followed by DP-16QAM misclassified as DP-QPSK and DP64QAM. Likewise, the strongest contributor for reducing classifier

1×5 neurons, Average accuracy 92.69%

a MLP / true	DP-BPSK	DP-QPSK	DP-16QAM	DP-64QAM	Accuracy
DP-BPSK	571	82	0	0	87.44%
DP-QPSK	32	10035	237	0	97.38%
DP-16QAM	1	286	6153	272	91.66%
DP-64QAM	0	1	132	2195	94.3%

1×10 neurons, Average accuracy 93.46%

b MLP / true	DP-BPSK	DP-QPSK	DP-16QAM	DP-64QAM	Accuracy
DP-BPSK	567	86	0	0	86.8%
DP-QPSK	23	10148	133	1	98.47%
DP-16QAM	0	210	6368	133	94.87%
DP-64QAM	0	0	145	2182	93.7%

1×100 neurons, Average accuracy 97.42%

c MLP / true	DP-BPSK	DP-QPSK	DP-16QAM	DP-64QAM	Accuracy
DP-BPSK	626	27	0	0	95.8%
DP-QPSK	21	10210	73	0	99.08%
DP-16QAM	0	120	6475	117	96.46%
DP-64QAM	0	0	38	2289	98.34%

2×100×10 neurons, Average accuracy 98.39%

d MLP / true	DP-BPSK	DP-QPSK	DP-16QAM	DP-64QAM	Accuracy
DP-BPSK	642	11	0	0	98.3%
DP-QPSK	31	10197	76	0	98.96%
DP-16QAM	0	71	6560	81	97.73%
DP-64QAM	0	0	33	2294	98.57%

Figure 4. Confusion matrices for various ANN architectures: (a) 1 hidden layer 5 neurons, (b) 1 hidden layer 10 neurons, (c) 1 hidden layer 100 neurons, and (d) 2 hidden layers 100 and 10 neurons.

accuracy for MLP_B and MLP_C is also misclassification of DP-BPSK to DP-QPSK, except for MLP_D, where DP-16QAM is misclassified as DP-QPSK and DP-64QAM. The key error contribution may be attributed to very similar cardinality properties for DP-BPSK and DP-QPSK.

4. CONCLUSIONS

Modulation format prediction based on current network state information is a promising approach to improve physical layer resource utilization, and near performance-limit operation over channel lifetime. We presented, for the first time to our knowledge, several MLP classifiers to predict physical layer modulation formats based on a comprehensive dataset. Specifically, our weakest classifier achieved an accuracy of ~92.7% with corresponding model training time of 75 s, whereas our strongest 2-layer MLP classifier achieved accuracy of ~98.4% with model training time of ~20 s. Furthermore, we ascertained that the strongest misclassification contributor for most MLP based models was DP-BPSK being miscategorised as DP-QPSK.

Our preliminary results suggest that ML could prove to be a valuable tool for real-time data-driven transmission capacity optimization, offering a dynamic and scalable solution to next-generation network resource optimization. For future work, we intend to compare MLP based models to other ML methods, together with dataset enrichment through in-field data.

ACKNOWLEDGEMENTS

We would like to acknowledge SPEED (BMBF FKZ 13N13744), SENDATE SECURE-DCI (C2015/3-4), and METRO-HAUL (G.A 761727) projects for partially funding this work.

REFERENCES

[1] S. Shahkarami *et al.*: Machine-learning-based soft-failure detection and identification in optical networks, in *Proc. OFC*, 2018, paper M3A.5.
[2] Y. Huang *et al.*: Dynamic mitigation of EDFA power excursions with machine learning, *Opt. Express* 25, 2245-2258, 2017.
[3] D. Rafique *et al.*: Performance improvement by fibre nonlinearity compensation in 112 Gb/s PM M-ary QAM, in *OFC*, 2011, paper OWO6.
[4] G. Khanna *et al.*: Comparison of single carrier 200G 4QAM, 8QAM and 16QAM in a WDM field trial demonstration over 612 km SSMF, in *Proc. ECOC*, 2016, paper W3D.3.
[5] D. Rafique *et al.*: Enabling 64Gbaud coherent optical transceivers, in *Proc. OFC*, 2017, paper W2A.25.
[6] D. Rafique *et al.*: Multi-flex field trial over 762km of G.652 SSMF using programmable modulation formats up to 64QAM, in *Proc. OFC*, 2016, paper W4G.2.
[7] R. Casellas *et al.*: Control, management and orchestration of optical networks: Evolution, trends and challenges, in *Proc. ECOC*, 2017, paper Tu1G1.
[8] M. Channegowda *et al.*: Software-defined optical networks technology and infrastructure: Enabling software defined optical network operations [Invited], *JOCN* 5, 10, pp. A274-A282, 2013.
[9] S. Yan *et al.*: Field trial of machine-learning-assisted and SDN-based network planning with network planning database, in *Proc. ECOC*, 2017, paper ThPDPB4.
[10] T. Gui *et al.*: Opportunities for machine learning in optical communication systems and networks, in *Proc. ACPC*, 2017, paper Su4B.1.
[11] L. Velasco *et al.*: First experimental demonstration of autonomic slice networking, in *Proc. OFC*, 2017, paper Th5A3.
[12] D. Rafique *et al.*: Analytics-driven fault discovery and diagnosis for cognitive root cause analysis, in *Proc. OFC*, 2018, paper W4F.6.
[13] A.P. Vela *et al.*: Combining a machine learning and optimization for early pre-FEC BER degradation to meet committed QoS, in *Proc. ICTON*, 2017, paper We.B3.1.
[14] X. Chen *et al.*: Leveraging deep learning to achieve knowledge-based autonomous service provisioning in broker-based multi-domain SD-EONs with proactive and intelligent predictions of multi-domain traffic, in *Proc. ECOC*, 2017, paper W.3.A.3.
[15] D. Rafique *et al.*: TSDN-enabled network assurance: A cognitive fault detection architecture, in *Proc.* ECOC, 2017, paper W3A4.
[16] P. Samadi *et al.*: Quality of transmission prediction with machine learning for dynamic operation of optical WDM networks, in *Proc. ECOC*, 2017, paper W.3.A.1.
[17] L. Barletta *et al.*, QoT estimation for unestablished lightpaths using ML, in *Proc. OFC*, 2017, paper Th1.J.1F.6.
[18] P. Poggiolini *et al.*: Recent advances in the modeling of the impact of nonlinear fiber propagation effects on uncompensated coherent transmission systems, *JLT* 35, 3, 458-480, 2017.