Using Multi-Layer Perceptron and Complex Network Metrics to Estimate the Performance of Optical Networks

Danilo R. B. de Araújo, Joaquim F. Martins-Filho
Department of Electronics and Systems
Federal University of Pernambuco
Recife, 50.740-550, Pernambuco, Brazil
E-mail: jfmf@ufpe.br

Carmelo J. A. Bastos-Filho
Polytechnic School of Pernambuco
University of Pernambuco
Recife, 50.720-001, Pernambuco, Brazil
E-mail: carmelofilho@ieee.org

Abstract—The performance assessment of a WDM network considering physical impairments is a difficult task and is frequently accomplished by using time consuming computational simulations. On the other hand, we observed that several metrics have been proposed to assess different aspects of a network structure. In this paper we propose that a set of metrics can be combined in order to obtain a fast estimation of a WDM network performance, based on a historical database of networks. The estimator was obtained by means of the most used Artificial Neural Network (ANN) architecture, called Multi-Layer Perceptron, that was trained using the classical back-propagation algorithm. According to our results, it is possible to build an estimator based on network metrics that assess WDM networks considering the trade-off between the processing time and the precision of the results. Our study also suggests that this kind of estimator can be easily adapted to other scenarios of WDM networks since Artificial Neural Networks present interesting characteristics, such as adaptation and flexibility.

Keywords—Optical Networks, Network Assessment, Complex Networks, Artificial Neural Networks, Multi-Layer Perceptron.

I. INTRODUCTION

Optical communications systems are evolving to wavelength-routed optical networks, which can provide reconfigurable high-speed circuit-switched connections between edge optical line terminals, named lightpaths. In general, the lightpaths are routed through core elements, such as optical add/drop multiplexers (OADMs) and optical cross-connectors (OXC).

The analysis of the propagation and degradation of a signal over a optical fiber link is a task that has been well aided in recent years by software tools. Generally, these tools use analytical models or simulations of wave propagation, using Split-Step Fourier method for example, which are accurate but time consuming. As an example, a WDM signal propagation in a single fiber link may take minutes to be evaluated.

Considering the emergence of wavelength-routed optical network, we also expect to find suitable tools for the analysis of high-level functions in optical networks. However, moving from the link to network analysis increase the complexity of the computational tool in the same proportion of the problem complexity. In network analysis, the use of wave propagation method is computationally too expensive and is often prohibitive. Therefore, network analysis tools often use approximate analytical expressions that attempt to model the behavior of the optical signal in a network environment. The analytical expressions for each physical impairment can be used to obtain a single metrics to assess the whole network performance. The blocking probability (BP) is an example of metrics to assess the network performance and it gives the probability of a connection request to be denied. The BP is usually estimated in analysis tools by using the Monte Carlo method. This kind of analysis tool generates a large number of request calls, evaluates the admission of these calls in the networks and obtains the BP by dividing the number of blocked calls by the total number of calls. Chaves et al. developed the SIMTON, a simulation tool for transparent optical networks [1]. SINTOM uses the BP to assess network performance. In [1] we can notice that it is necessary $1.5 \cdot 10^5$ request calls to evaluate a network with a BP level at $8 \cdot 10^{-3}$. According to Chaves et al., if all the physical impairments are considered by the SIMTON, the total time to evaluate the Finland network is up to 1.25 seconds.

Network analysis tools can provide information of optical network performance based on the overall network environment, namely: the physical topology; the configuration of the allocated devices in optical terminals; number of wavelengths; and the specifications of traffic demands. If the specification of traffic is kept fixed, it is possible to associate the performance of the network with the physical topology and with the other variables that could change the signal propagation. The task of trying to capture complex relationships of the topological structure with optical devices can be a more high level approach to an analysis tool for optical networks. This kind of tool would provide support for stake-holders to move from the analysis of small ring networks to the analysis of optical mesh networks with hundreds or thousands of nodes.

Complex Networks or Network Science is a field of study that is becoming popular mainly due to their capability to represent virtually any network structure of real-world phenomena. Because of this, several investigations have been proposed to represent the network structure and to analyze the topological features of the networks in terms of Network Metrics (NM), including the analysis of dynamical changes on the topology over the time. There are some recent studies focusing on the relationship between the structure and the
dynamics of Complex Networks [2]. NM are quite important to obtain a comprehensive understanding about this relationship. Besides, the quantitative description of the networks properties also provides fundamental subsidies for classifying and understanding Complex Networks.

This paper aims to present a fast method to assess the performance of optical networks based on NM. Our main goal is to develop a simple method to estimate BP using Artificial Neural Networks, instead of using the time consuming Monte-Carlo computational simulations. The remainder of this paper is organized as follows: Section II provides a brief review on the most used NM to characterize Complex Networks; Section III presents our proposal based on Artificial Neural Networks to estimate the performance of optical networks; Section IV and V present the simulation setup and some preliminary results; and Section VI presents the conclusions and future works.

II. A BRIEF SURVEY ON COMPLEX NETWORK METRICS

Relevant topological properties of a network are highly correlated to the graph that represent this network. These properties are often used to define a family of graphs and, because of this, the definition of this type of correlation is one of the major research lines on Complex Networks. Some NM have been proposed to quantify relevant topological properties of Complex Networks [3]. This section presents a brief review of the most used NM [4], [5], [6], [7].

In this paper, we consider a Complex Network as an unweighted and undirected graph \( G = (N, L) \), in which \( N \) and \( L \) denote the set of vertices and the set of edges, respectively. The amount of nodes and links in a network are defined as \( N = |N| \) and \( L = |L| \), respectively. Besides, we consider that a graph cannot contain connections beginning and ending at the same node.

The most used topologic NM are presented in this paragraph. The link density \( (q) \) of a network is defined as the ratio between the number of links that actually exists and the number of links if all nodes are connected to all others by direct links. The node degree \( (D) \) describes the number of links or neighbor nodes of a given node. The node degree distribution defines the probability, \( Pr(D) \), of a randomly selected node to have a certain degree \( D \). The average number of links that are connected to a node is called the average node degree. The entropy provides an average measurement of the heterogeneity of the network. The shortest path describes the minimum number of hops between a given pair of nodes. The distance distribution is the probability, \( Pr(H) \), that a randomly selected pair of nodes presents a shortest path with value equal to \( H \) hops. The longest shortest path \( (H_{max}) \) between any pair of nodes is referred as the diameter of a graph, which is defined as \( diam(G) \). The Average Path Length is the average number of hops of the shortest paths considering all the source-destination pairs. The clustering coefficient \( (c) \) is the ratio between the number of triangles that contain node \( i \) and the number of triangles that could possibly exist if all neighbors of \( i \) were interconnected [8]. The clustering coefficient for the entire graph \( (c_G) \) is the average of the clustering coefficients of all the network nodes. A graph is said to be connected if there exists a path between each pair of nodes. When there is no path between at least one pair of nodes, a network is defined as disconnected. If there is a link between every pair of nodes in a graph, the graph is defined as complete \( (K_N) \). The link connectivity \( k_\ell \) is the minimum number of links to be removed in order to turn a graph disconnected. The node connectivity \( k_N \) is defined analogously [9].

In the graph theory, a network can be represented by its Adjacency matrix \( (A) \), the Node Degree matrix \( (D) \) or the Laplacian matrix \( (L) \). A is a \( N \times N \) matrix, in which the non-diagonal entries \( (i, j) \) are equal to “1” if the nodes \( i \) and \( j \) are connected, or “0” otherwise. In \( A \), the entries \( (i, i) \) are always equal to “0”, since we are not considering self-loops. Since we are considering undirected graphs, \( A \) is a symmetric matrix. \( D \) is diagonal matrix which contains information about the degrees of the nodes. \( L \) is defined as \( L = D − A \), in which the non-diagonal entries \( (i, j) \) are either “−1” or “0”, depending on whether nodes \( i \) and \( j \) are connected or not, respectively, and the diagonal entries \( (i, i) \) are equal to the degree of the nodes \( D_i \). Some NM are closely related to the eigenvalues of the matrices that represent the network. These NM are known as spectral metrics. All eigenvalues are real for \( A \) [10], whereas all eigenvalues are real and nonnegative for the \( L \) [11]. The ordered set of \( N \) eigenvalues of \( A \) or \( L \) is called the spectrum of the matrix and can be used to classify the network generational model. The largest eigenvalue of \( A \) is denoted as the spectral radius \( (\rho) \), whereas the second smallest eigenvalue of \( L \) is denoted as the algebraic connectivity \( (\lambda_{N−1}) \). A graph is disconnected if \( \lambda_{N−1} = 0 \). Moreover, if \( \lambda_{N−1} = 0 \) and \( \lambda_{N−2} \neq 0 \), then a graph has exactly \( i \) components. This also means that the multiplicity of zeros in the eigenvalues of the Laplacian matrix corresponds to the number of independent components of the graph. \( \lambda_{N−1} \) also measures the connectivity of a graph, i.e. a higher value for \( \lambda_{N−1} \) implies in a higher difficulty to cut a graph in two independent components [12]. Another important spectral metric, called natural connectivity NC, aims to characterize the redundancy of alternative routes in a network. NC is calculated by quantifying the number of closed chains of all lengths in the network [13]. Both \( \lambda_{N−1} \) and NC are commonly used to measure the robustness of real-world networks.

III. OUR PROPOSAL

The main hypothesis of this work is that a very complex attribute of a WDM optical network, such as the BP, could be estimated by using a combination of some NM and some general characteristics of the physical layer. First, we selected the most used NM and we performed a cross-correlation analysis among these NM. The We have observed from a preliminary analysis that only six of the studied NM do not present high correlation when performing network performance assessment. Thus, we selected the following metrics for the regression: algebraic connectivity, natural connectivity, average path length, clustering coefficient, diameter and entropy. We also included two characteristics of the physical layer, the OXC type and the maximum number of wavelengths. We chose these two characteristics because the first is related to the nodes and second is related to the links. Artificial Neural Network (ANN) is a well known tool for modeling systems that present complex and nonlinear relationships between inputs and outputs [14]. Multi-Layer
Perceptron (MLP) is the most used ANN architecture and can be used for classification, forecasting or regression tasks. Each artificial neuron is specified by a set of weights, a bias input and an activation function. We used the sigmoid logistic function as the activation function [14]. Since a MLP with one hidden layer is simple and can tackle nonlinear problems, we used this topology to forecast the PB. Figure 1 presents our proposal with eight inputs in the first layer, in which six of them are NM. The output layer has one artificial neuron and the output of this neuron gives the estimated blocking probability.

![Diagram of the proposed architecture to the Artificial Neural Network.](image)

**Fig. 1**: The proposed architecture to the Artificial Neural Network.

### IV. Simulation Setup

This section aims to provide the information for training and evaluation of the MLP. Besides, we briefly describe the setup for the Monte-Carlo simulation of the optical networks to determine the reference BP values.

#### A. General Physics Layer Characteristics for the MLP

We considered 5 different types of OXCs. The available devices are presented in Table I. In this case, different OXCs leads to different BPs because of the crosstalk between switch ports. The labels were normalized to values in the interval $[0, 1]$ to be used as inputs of the MLP. Each WDM network could have at most 40 wavelengths. The number of wavelengths used as input to the ANN is also normalized to the interval $[0, 1]$.

<table>
<thead>
<tr>
<th>Label</th>
<th>Normalized label</th>
<th>Isolation Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>-30 dB</td>
</tr>
<tr>
<td>2</td>
<td>0.25</td>
<td>-33 dB</td>
</tr>
<tr>
<td>3</td>
<td>0.5</td>
<td>-35 dB</td>
</tr>
<tr>
<td>4</td>
<td>0.75</td>
<td>-38 dB</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>-40 dB</td>
</tr>
</tbody>
</table>

**TABLE I**: Labels and specifications for the available OXCs.

#### B. Training and Evaluating the MLP

We used 32,000 different network topologies obtained from the same node distribution of the NSFNET benchmark optical network to evaluate our proposal. The entire set of networks was divided in three different groups to be used for distinct goals: training, validation and test. We used 16,000 training samples and 8,000 validation samples during the training process. We organized the sets in order to have samples for the entire PB range in all sets. The training set is used to adjust the weights of the MLP. We used the validation set to check if the MLP is not overfitting the samples from the training set. The training must stop when the validation error (calculated by using the validation set) starts to increase or when the number of iterations reach $I_{\text{max}} = 100,000$ [14]. In each trial, the trained MLP was evaluated in the set of test samples with $n = 8,000$ optical networks. The entire set of optical networks was evaluated by the SIMTON (see next subsection) in order to obtain the BP values to be used as the reference for the output.

We used the mean squared error (MSE) as the metrics to assess the performance of our proposal. $MSE$ is defined in Eq. (1).

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (BP_{\text{SIMTON}}, - BP_{\text{ANN}})^2.$$  \hspace{1cm} (1)

where $n$ is the number of patterns, $BP_{\text{SIMTON}}$ is the BP obtained by the SIMTON simulator and $BP_{\text{ANN}}$ is the BP estimated by the MLP.

We used the sigmoid logistic function with parameter randomly selected in the interval $[0, 1]$. We used the back-propagation algorithm for training with an initial learning rate equal to 0.05 [14]. The learning rate is multiplied by $0.5$ in each iteration, where $I_{\text{max}}$ is the maximum number of iterations. The bias of the neurons are initialized randomly in the interval $[0, 1]$. The biases are updated by the back-propagation algorithm during the training and can assume values in the interval $[0, 1]$. All NM are set to vary in the interval $[0, 1]$. The number of neurons in the hidden layer ($N$) is a parameter of the proposal and is analyzed in this paper.

#### C. Physical Layer Model and Parameters used to Obtain the Reference Blocking Probabilities

We used the SIMTON ( Simulator of Transparent Optical Networks) [1] to estimate the BP for each one of the 32,000 optical network samples. The Optical Signal to Noise Ratio (OSNR) of each lightpath is evaluated considering the following impairments: the amplified spontaneous emission (ASE) noise of the amplifiers, the saturation effect of the amplifier gain, the saturation of the ASE noise in EDFAs, the residual chromatic dispersion, the crosstalk in the OXC and the Four Wave Mixing (FWM) effect. The values for the network parameters used in the simulations are shown in Table II.

### V. Results

We run 30 trials for four different numbers of neurons in the hidden layer, $N = 5, 10, 15$ and 20. The box plots for the $MSE$ are presented in Figure 2. One can observe that the MLP with 5 neurons in the hidden layer ($4N^2/5$) presents an outlier. For the back-propagation training algorithm, there is no significant difference between $N = 10, 15$ and 20 neurons in the hidden layer.
Figure 3 shows the evolution of the training MSE and validation MSE for one execution trial with $N = 15$. One can observe that both MSE quickly decay to a low level and stagnates after approximately 10,000 iterations.

For practical purposes, it is very important to analyze the performance of our proposal in term the MSE for different intervals of the BP. In general, WDM optical networks present BP below 1%. In order to allow this analysis, we divided the test set in four groups: Data-set A, Data-set B, Data-set C, Data-set D represent sets of samples with BP in the intervals $(0.0; 0.25)$, $(0.25; 0.50)$, $(0.50; 0.75)$ and $(0.75; 1.0)$, respectively. Figure 4 presents the boxplot of the MSE for a MLP with $N = 15$. One can observe higher MSE values for Data-set A and Data-set B. This means that the MLP presents a better performance for higher BP values.

We also measured the execution time for both approaches in other to compare their performance. We executed $1.5 \cdot 10^5$ calls in the SIMTON, and the measured execution time was $1250 \text{ ms}$. On the other hand, once the MLP is trained, the execution time to estimate the BP is below $1 \text{ ms}$ for our approach.
VI. CONCLUSION

In this paper we presented an alternative way to assess the performance of WDM networks. Our method consists in mapping complex network metrics to $BP$ values in order to estimate the $BP$ of an optical network. This kind of mapping is taken place by using a MLP neural network. We used eight different attributes as an input to the neural network and we used one hidden layer with several configurations to combine the inputs. According to our results, it is possible to build a simple estimator with reasonable precision for estimating the $BP$ of WDM networks.

Further analysis aims to investigate if different new metrics in the input could offer a better precision of the output metrics. Also, we recommend the use of a more sophisticated training algorithms based on global optimizers, such as particle swarm optimization, differential evolution or genetic algorithms. Another analysis could focus on the self-adjust capability of the neural network in order to include other optical network scenarios, such as different traffic distributions and variations in the network topology.

We also suggest that our proposal could be used together with a more precise $BP$ estimator in order to obtain a desired trade-off between high precision and low time processing. We also would like to emphasize the difference between the total execution time related to the two methods. The obtained speed-up motivates the use of our proposal as an alternative way to evaluate the performance of WDM networks, specially in the cases where too many Monte Carlo simulations are necessary to evaluate the $BP$.

ACKNOWLEDGMENT

The authors acknowledge the financial support from FACEPE, CNPq, CAPES, UFPE and UPE.

REFERENCES