Machine Learning Assisted Traffic-Aware Approach to Path Assignment in SDM-EONs

Ramon A. Oliveira[1], Denis Rosário[1], Eduardo Cerqueira[1] and Helder Oliveira[1]

[1]Federal University of Pará (UFPA)
Rua Augusto Correa, 01, Guamá – Belém – PA – Brazil

Abstract. *The introduction of new technologies and applications connected to the Internet has demonstrated the inability of current optical networks to provide resources for next-generation Internet. Although the emergence of elastic optical networks with space-division multiplexing has shown to be a promising solution to deal with the capacity problem, some of the technical requirements for the implementation of these networks remain open challenges. In this sense, this paper proposes MISSION, a Machine Learning assisted, fragmentation, and crosstalk-aware model for path allocation in Space Division Multiplexing Elastic Optical Networks (SDM-EONs). The proposed approach is capable of ordering candidate paths for allocation based on metrics such as crosstalk, fragmentation, and the number of slots. Besides, MISSION shows competitive performance, by keeping a comparatively low blocking probability and fragmentation, even under heavy loads.*

1. Introduction

The ever-rising adoption of cloud computing, on-demand video, Internet of Things and so many other emerging services have increased traffic and are leading to the exhaustion of the current transmission capacity of the internet [Cisco 2018]. Optical fiber-based networks are an essential part of this scenario, responsible for connecting billions of users to the Internet in the whole world. As new applications drive the need for larger bandwidths, new and more reliable connections become necessary. However, the mono-mode fiber infrastructure broadly adopted in current-gen optical fiber networks has a limited physical transmission capacity unlikely to be improved soon [Ítalo Brasileiro et al. 2020]. According to some estimates, throughput available to the end-user should grow from 100 Mb/s to 1 Gb/s in the next ten years, which would force the network's core to shift to 1 Tb/s, and require much more resources for data transfer [Ítalo Brasileiro et al. 2020].

Recently, Space-Division Multiplexing Elastic Optical Networks (SDM-EONs) have drawn the attention of the scientific community, mainly for being a promising way to increase transmission capacity in optical fiber networks, as well as for providing flexibility in resource allocation. SDM-EONs was developed as an improvement upon traditional optical network infrastructures that increase the available resources and allow for a more efficient mechanism for their allocation [Oliveira and Fonseca 2019] by utilizing multiple spatial channels for data transmission. In an SDM-EON network, the Routing, core, modulation level, and spectrum assignment (RCMLSA) problem refer to a set of challenges to the allocation of network resources. The RCMLSA is a process that allows for the establishment of reliable transmission channels in a given network by determining core, modulation, and path selection.

In SDM-EON, for a connection to be successfully established the fundamental continuity and contiguity constraints must also be strictly followed. These constraints ensure that transmissions occur on a single core along the entire route and that the slots allocated for the request must be contiguous [Zhu et al. 2021]. Analytic approaches to various problems pertaining to RCMLSA in SDM-EONs are well documented in the literature and the heuristics developed were shown to perform well [Beyragh et al. 2019], [Paira et al. 2020] and [Yousefi and Rahbar 2020]. However, some disadvantages of such approaches are the labor-intensive development, the lack of adaptability of the developed algorithms, as well as the waste of diagnostic, statistic and historical data continuously produced during the operation of such networks that goes unused.

In this context, Machine Learning (ML) appears to provide adaptability to path assignment in the SDM-EON context. An ML model is a statistic model applied to a complex set of heterogeneous data so that it reveals patterns in the data that can be used to classify or categorize unknown data that share the feature space. ML can be further divided according to the organization of the employed data. Unsupervised ML utilizes non-labeled data that can be clustered or associated. On the other hand, supervised ML consists of using labeled data, whose labels can be used for classification or regression tasks [El Naqa and Murphy 2015]. This paper proposes an approach ***M**achine Learn**I**ng A**SS**isted Traff**IC**-Aware Appr**O**ach to Path Assig**N**ment in SDM-EONS*, called MISSION. MISSION employs both classification and regression for path allocation optimization in SDM-EONs. MISSION aims at reducing the crosstalk and fragmentation levels acceptable under different loads.

MISSION considers a dataset consisting of traffic and diagnostic data of a simulated SDM-EON comprising data on crosstalk, fragmentation, number of slots and path ID. This data is used to train an ML model to estimate the probability of any given path being accepted for allocation, given the current network conditions as described by the aforementioned features. This novel acceptance metric is then applied to sort through a set of possible paths, according to how likely they are to being successfully accepted. The MISSION is shown to be over 97% accurate in predicting whether a path would be successfully accepted and network simulations performed in a SDM-EON that incorporates our approach showed that it maintains competitive crosstalk and fragmentation levels, while keeping low blocking probability, compared to other similar approaches from the literature.

The remainder of this paper will be organized as follows: Section 2 presents the related works to the researched subject. Section 3 presents a detailed description of the MISSION approach, a detailed description of the network abstractions used for data-gathering and testing, the composition of the dataset used for the ML-model training, and the ML-algorithm adopted for the final model. Section 4 presents the comparative results of our approach and a detailed account of the simulation parameters adopted. Section 5 discusses future developments to the proposed approach.

2. Related Works

[Trindade and da Fonseca 2019] proposed two proactive algorithms to avoid spectrum fragmentation that take into account the fragmentation state and possible formation of bottlenecks in the network. The first uses, along with the fragmentation metric, a closeness centrality of a node metric to avoid the allocation of the spectrum to adjacent cores, reducing inter-core crosstalk. The second algorithm uses a pre-determined order for core allocation upon which a first-fit policy is applied in order to attempt allocation.

[Trindade and da Fonseca 2021] introduced two algorithms, one proactive and the other reactive, combined to an unsupervised ML algorithm that form an approach to RM-CSA in which the clustering of lighpath data is used to establish lighpaths for incoming calls. This approach was shown to reduce spectrum fragmentation, as well as the blocking of requests. Ours uses many of the aforementioned traffic data, although the ML-approach used in our approach is supervised.

[Zhao et al. 2018] proposed a crosstalk-aware, spectrum defragmentation approach that moves connections in order to achieve higher spectrum compactness. Such approach introduces a novel metric that describes spectrum defragmentation in each core. The researchers developed two defragmentation solutions, namely, Same Spectrums and Different Cores (SS-DC) and Different Spectrums and Same Cores (DS-SC). The former moves a connection to another core in the same link in the same spectrum slots, thus increasing slot compactness, and the later builds upon the SS-DC and, if either the crosstalk threshold is surpassed, or if no spectrum resources are available on other cores, the connection is to be moved to other available spectrum on the same core along the lightpath. By combining both strategies the researchers achieved better bandwidth blocking probability, spectrum utilization, spectrum moving times and spectrum defragmentation latency than the algorithm used for benchmarking.

[Yao et al. 2018b] introduced an approach in which crosstalk estimation is done by an ML-assisted model, to achieve higher accuracy. The reserchers adopted three ML algorithms, namely, Levenberg-Marquardt (LM), Bayesian Regularization (BR) and the Scaled Conjugate Gradient (SCG), of which the LM performed best. The resulting crosstalk estimation data is then used by a core, mode and spectrum assignment algorithm that uses a lower frequencies-first allocation policy, which is the most efficient, according to the researchers.

[Yousefi et al. 2020] introduced two new metrics employed in three novel algorithms aimed at improving blocking probability in SDM-EONs by reducing fragmentation. The coefficient of variant metric (CVM) calculates fragmentation as a function that incorporates the area occupied by the connection, the number of adjacent connections within a certain distance from the allocated connection and the total number of connections within the allocated rectangular region. The holding time metric (HTM) calculates the holding time of the connections, so that it can be employed to the allocation of connections near connections that will almost immediately be released, thus reducing fragmentation. The researchers found that the algorithm that employs the second metric was successful in reducing blocking probability, reducing spectrum utilization and in reducing fragmentation.

[Yang et al. 2021] employed a multi-input multi-output multi-task convolutional neural network composed of multiple input/output layers for multi-task optical performance monitoring in SDM-EONs. The results showed a 100% accuracy for modulation format identification for all signals. Optical signal-to-noise ratio could also be estimated within a 0.6dB margin of error. The resulting model uses a reduced dataset, has a lower computational and represents an easily implementable mechanism for network optimization.

[Zhu et al. 2021] proposes a Fluctuating Traffic Model that describes fluctuating traffic-varying flow rates, and a Protected Routing, Modulation, Spectrum and Core Allocation (RMSCA) algorithm that minimizes crosstalk and blocking probability at peak loads, by utilizing a triangular iterative core assignment strategy. Unlike the MISSION algorithm, the authors attempt to reduce XT specially at peak loads of fluctuating traffic and employ an entirely heuristic approach to achieve their positive comparative results.

[Yao et al. 2018a] addressed the spectrum optimization problem by employing transductive transfer learning, which allows for predictive models to be built even if the training and test datasets come from different feature spaces. This ML approach was employed in the proposed model to predict the spectrum migration time used for resource reservation capable of reducing blocking probability, thus increasing spectrum utilization. Unlike the study reviewed above, ours attempts to build a broader dataset, that includes not only blocked services, as well as fragmentation, crosstalk, and path statistics in order to arrive at a more reliable model.

[Moura and da Fonseca 2018] proposed a novel RCMLSA approach to search for available spectrum in SDM-EONs by means of binary image processing algorithms. From this unified approach four algorithms can be derived, according to the fitting policy adopted, Connected Component Labeling-Best-Fit (CCL-BF),CCL-Random-Fit (CCL-RF), Inscribed Rectangles Algorithm-Minimal-Blocking (IRA- MB), and IRA-Minimal-Crosstalk (IRA-MC). Our simulations adopt this approach for data collection and as a baseline for testing the performance of our ML model.

[Rodrigues et al. 2021] presented an RMSCA heuristic for blocking mitigation in SDM-EON. This approach calculates the average crosstalk and fragmentation for a given number of k-nearest paths before allocation and prioritizes the path with with the lowest predicted value in those metrics for allocation. Our simulations also adopt this approach for benchmarking purposes, although no simulation data from this algorithm was employed in our model's training.

The analyzed works show several approaches for SDM-EON optimization, such as the estimation of key metrics to enhance performance, the creation of new metrics that allow for better Quality of Transmission (QoT) estimation, ML-aided lightpath establishment for incoming calls, as well as accurate modulation estimation. Table 1 presents a comparison between some related work and the proposed algorithm in this paper. The works were categorized with respect to the use of SDM, use of ML, and, finally, whether they presented and employed a novel metric (METRIC) for network optimization. In this paper, we have likewise relied on a structured and selected dataset of various key network aspects. Considered aspects are more diverse than those present in the literature, and made it possible to build a highly reliable, easily implementable ML model for path assignment in SDM-EONs. Besides that, due to its diversity, our model ML shows capable of rivaling current competing approaches to the researched problem.

Table 1. Comparison of related works

Approach	SDM	ML	METRIC
[Trindade and da Fonseca 2019]	✓	✓	✓
[Trindade and da Fonseca 2021]	✓	✓	
[Zhao et al. 2018]	✓		✓
[Yao et al. 2018b]	✓	✓	
[Yousefi et al. 2020]	✓		✓
[Yang et al. 2021]	✓	✓	
[Zhu et al. 2021]	✓		
[Yao et al. 2018a]	✓	✓	
[Moura and da Fonseca 2018]	✓		
[Rodrigues et al. 2021]	✓		
MISSION	✓	✓	✓

3. MISSION Algorithm

This section introduces the MISSION approach, which employs an ML predictive model trained with synthetic traffic data that analyses candidate paths to be assigned to a particular connection and estimates how likely the path is to be accepted. Afterwards, it sorts those paths accordingly, which leads to a path assignment approach that relies not only on the distance but is also influenced by past network data of key performance metrics. The following subsections further elaborate on the notation used, the data collection stage, feature extraction stage, training stage and algorithm selection stage that compose the proposed approach.

3.1. Notation

We employed the following notation for the network abstraction used in this paper:

s: source node;

d: destination node;

b: bandwidth demand in slots;

N: number of the slot between two nodes;

C: number of cores;

V: set of nodes;

$E = \{e_{u,v,n}\}$: set of edges connecting u and v;

$G = (V, E, W)$: labeled multigraph composed by a set of nodes V, a set of edges E and a set of edge weights W, $|E| = C \cdot N \cdot |V|$. The edges connecting two vertices of G represent the N slots in the link connecting two network nodes;

S_{ij}: occupancy matrix representing the spectrum of the link between u and v in G.

3.2. Network Overview

The optical network employs spatially flexible reprogrammable optical add/drop multiplexers that allow wavelength-selective switch, and space-wavelength granularity, with multiple-input multiple-output (MIMO) transceivers. The network complies with the fundamental continuity and contiguity constraints, *i.e.*, , the allocation was performed in the same spectrum in each fiber along the route of a lightpath, and slots were allocated continuously in the spectrum. Figure 1 illustrates how these constraints are applied to the allocation process in the adopted network abstraction. Of the five links shown, all of which have available slots, only links 1, 3, 4 and 5 have at least two available slots that are both continuous and contiguous, *i.e.*, , slots available side-by-side in their respective links and in the same position relative to the remaining links so that data transmission from nodes A to B requiring two slots would have to go through links 1, 5, 4 and 3, even though the path that goes through link 2 also has at least two available slots and is of shorter length.

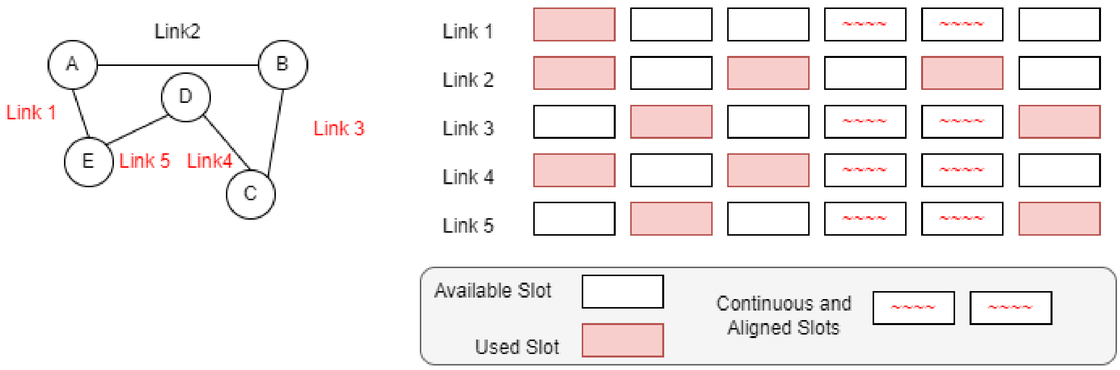


Figure 1. Continuity and Contiguity Constraints

Table 2 shows the modulation characteristics considered in our network scenario. One or more modulations might be available for a given request, according to the request's requirements and path configurations.

Table 2. Modulation Characteristics

Modulation Level	# Bits per Symbol	Slot Capacity (Gb/s)	Maximum Distance (km)
64QAM	6	75	125
32QAM	5	62.5	250
16QAM	4	50	500
8QAM	3	37.5	1000
QPSK	2	25	2000
BPSK	1	12.5	4000

3.3. Feature Extraction

The dataset employed to train the ML model used in this study was created with the Numpy[Harris et al. 2020] and pandas[pandas development team 2020] libraries and compiles data on the main performance indicators for SDM-EONS, sampled from a total of 180.000 attempts to allocate connections in a simulated SDM-EON topology. The data retrieved represented the *fragmentation*, *crosstalk*, *path ID*, *number of slots required*, and *call acceptance status*.

Numeric data (*fragmentation*, *crosstalk* and *number of slots*) was normalized so that it could be converted to meaningful input to the ML algorithm, while the remaining data received a binary mask since it cannot be expressed in quantities, as is the case with the combination of links that compose the prospective call path and the call acceptance status. By transforming such data in Boolean values, it can be processed by the ML algorithm without introducing bias.

Every path is considered for allocation, that is, every set of nodes V, connected by a set of edges E that describe a path between a source node s and a destination node d receives a unique identification, the *pathID*. Following the ID is the data on the network status when the allocation attempt was made and whether the allocation was successful or not. The complete dataset is structured as shown in Figure 2.

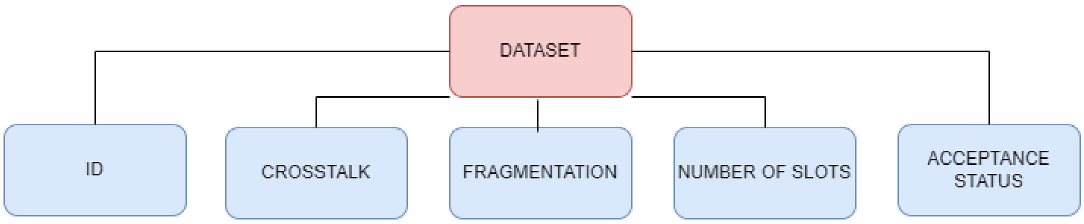


Figure 2. Dataset Structure

3.4. Training

The MISSION approach was trained with Scikit-learn's[Pedregosa et al. 2011] implementation of multiple ML algorithms, such as Random Forest, Support Vector Machine and the logistic function, among which the logistic function was shown to be the least computationally intensive and least complex, whilst showing very similar performance, being thus chosen for the final model.

The dataset built with data sampled during the feature extraction step described in Section 3.3 was divided into two parts, being 80% reserved for training and 20% reserved for testing. Figure 3 provides a simplified representation of the classification task modeling adopted for the training of the ML model that is part of the MISSION algorithm, in which the classifier attempts to predict whether a call would be accepted or denied (i.e. the *acceptance status*) based on the individual path it would take (*pathID*), the *crosstalk*, as measured between each connection attempt, the *fragmentation* levels in the network as a whole, as well as the required *number of slots*, resulting in a classification model that outputs results in the form of Boolean values of *True* or *False*, or 1 or 0, in the adopted notation.

A Grid Search trained the model multiple times using varying parameters and the parameters most successful in classifying the data were employed to tune the final version of the classifier. The classification was ultimately more than 97% accurate. The resulting model was then inserted in the simulation platform and directly employed to assist in path assignment. All network parameters were exactly the same for the baseline simulations and ML-assisted simulations.

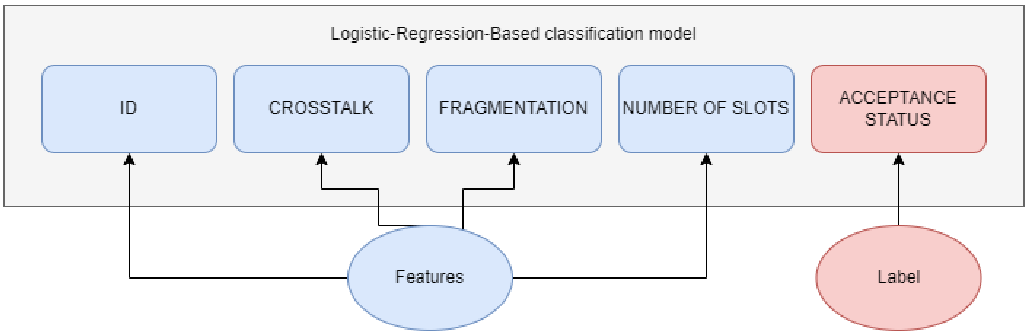


Figure 3. Model Training

3.5. MISSION Algorithm

The MISSION approach employs an ML model trained with synthetic network diagnostic data that analyzes candidate paths to be assigned to a particular call by how likely the path is to be accepted, and ranks those paths according to the estimated probability of acceptance, which leads to a path assignment approach that relies not only on the distance between nodes but is also influenced by past network data on key metrics. Our approach is built upon the Image-RCSA[Moura and da Fonseca 2018] algorithm, that utilizes binary image processing algorithms for the identification of available spectra and multiple fitting policies for resource allocation, depending of whether minimizing blocking or XT is more desirable.

To develop MISSION, the dataset built from network data was used to train several models based on the ML algorithms tested, being the logistic regression model the one selected, given its aforementioned qualities. Therefore, we employed the logistic regression model, as implemented in the Scikit-learn library. This model can be expressed by means of the logistic function, that will invariably result in values in the interval [0,1], so that it expresses the probability of a given event. This regression model allows for multiple features to be associated to a single dicotomic dependent variable.

$$f(z) = \frac{1}{1 + e^{-z}} \tag{1}$$

Logistic Regression based models for ML classification and regression tasks have been employed on SDM-EON research to accomplish many tasks in various of its subfields, such as QoT estimation [Khan et al. 2020], attack detection accuracy and others [Amirabadi 2019]. It was also shown in our experiments that, among some popular choices, such as Random Forest and Support Vector Machine, the Logistic Regression Algorithm was also the simplest to set and run a grid-search on (given the fewer tunable parameters) while maintaining virtually the same accuracy results as its rivals.

In order test the efficiency of the proposed ML model we proceeded to directly implement it in the simulation platform. The final ML model, more than 97% accurate, was exported using the Pickle[Van Rossum 2020] library and set up in a server by means of the Flask[Grinberg 2018] library. It was then included in the path allocation mechanism. To every connection attempt, a small dataset is formed, composed of n lines, equal to the number of candidate paths considered. Each line n is of the form *n* = *pathID*, *crosstalk*, *fragmentation, number of slots*.

Algorithm 1: MISSION

```
   Input: r(s, d, b)
   Output: List of Paths and their respective acceptance probability
1  P = kShortestPath(G,s,d,K);
2  for P_k ∈ P do
3      Gather current network data
4      Input data to predictive ML algorithm
5      output acceptance probability
6      Sort P according to acceptance metric
7  end
8  for P_k(sorted) ∈ P_sorted do
9      for m ∈ M do
10         for E_uv ∈ P_k(sorted) do
11             for S_ij do
12                 if Signal − to − Noise Ratio under threshold.
13                 & XTunderthreshold
14                 & S_ij = 0
15                 then
16                     Allocate call;
17                 else
18                     Block call;
19                 end
20             end
21         end
22     end
23 end
```

Algorithm 1 represents the MISSION approach and how the proposed ML-model works as part of the allocation process. The input to the model is a call requisition in which source and destination nodes are specified, as well as a demand in slots required for call allocation. The ML-assisted predictive model at the center of our approach then outputs, for every evaluated path, their respective probability of acceptance. Provided with the necessary data on the incoming call, the Dijkstra algorithm then calculates a list of five possible paths *P* (line 1) Then, every path P_k in set *P* is processed by the predictive ML model, as described by the loop in line 2.

Lines 3 through 6 show that, once a prospective path arrives, the algorithm performs the sampling of all relevant network data. This data includes the source and destination nodes and the links connecting them, that are converted into a unique ID. Then, it calculates the number of slots required, current network crosstalk and fragmentation as in [Moura and da Fonseca 2018]. Once all data is formatted, the resulting dataset becomes the input for the ML-model, which will process the data and output the path's acceptance probability, used to define how paths will be prioritized in the allocation process.

The following loop (lines 8 through 23) depicts the remainder of the allocation process. For every high-priority path (Line 8), in every modulation that can be employed for that path (Line 9), given the set of edges E_{uv} (Line 10) that connects source and destination, an occupancy matrix S_{ij} (Line 11) will be produced. By querying the occupancy matrix, the algorithm then verifies whether there are available continuous and contiguous slots and whether the signal-to-noise ratio and crosstalk are under pre-determined thresholds. If all conditions are true the connection is allocated, and, otherwise, the algorithm will attempt to allocate the next high-priority path.

4. Performance Evaluation

In this section, we elaborate on the performance tests of the MISSION algorithm. The model was implemented directly into the simulation platform and tested in the exact same configurations used for data collection and bench-marking, as described in Section 4.1. We chose to evaluate the main performance indicators for the intended purpose of this algorithm of mitigating blocked connections, i.e., the ratio of blocked calls to load (MBBR), fragmentation, and inter-core crosstalk (XT). To evaluate the model's performance, it was implemented and turned into an integral part of the network simulation environment, although all other configurations remained equal, which allowed for a fair comparison between our approach and the Image-RCSA[Moura and da Fonseca 2018] (IRA-MB variation), designed specifically for blocking optimization.

4.1. Simulation Description and Metrics

In this paper, we used the network simulator FlexGridSim [Moura and Drummond] to replicate an SDM-EON network operating on the NSF topology characterized by a seven-core fiber with 320-slot MCF links, each slot comprising 12.5 GHz, as depicted in Figure 4. In this topology, 14 nodes (numbered 0 through 13) represent the possible start or end-points for data transmission, connected bidirectionally by weighted vertices that describe the physical distance between the nodes it connects, as illustrated in Figure 5.

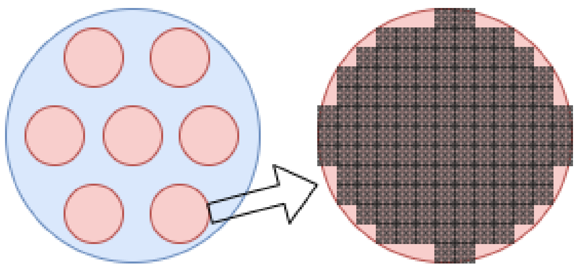


Figure 4. Core and Slot Distribution

Connection requests were modeled according to a Poisson process and are uniformly distributed among all node-pairs. The simulations were performed with loads between 100 and 500 erlangs, increased in each round of simulations by 50 erlang increments. The data generated by this simulation process was retrieved and further processed into a coherent dataset and used to train the ML model, as specified in the following sections. The metrics used to evaluate network performance were the medium bandwidth blocking ratio (MBBR), intercore crosstalk (XT), energy efficiency (EF) and the mean fragmentation as defined by Moura *et al.*[Moura and da Fonseca 2018].

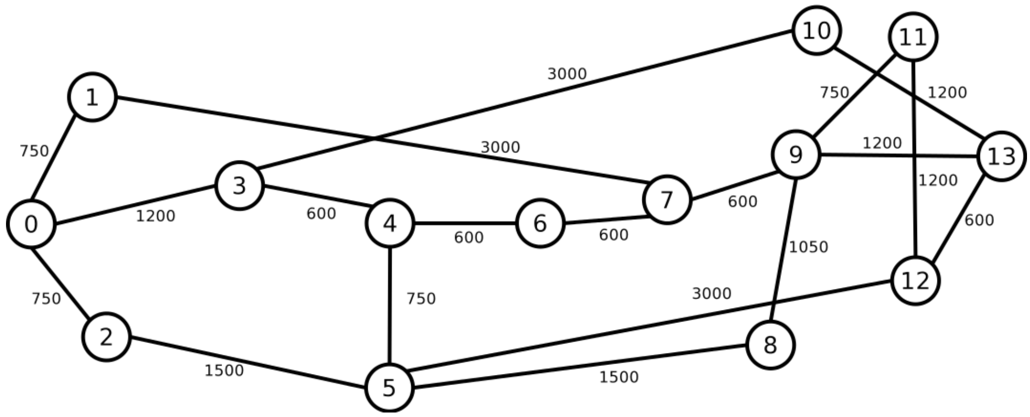
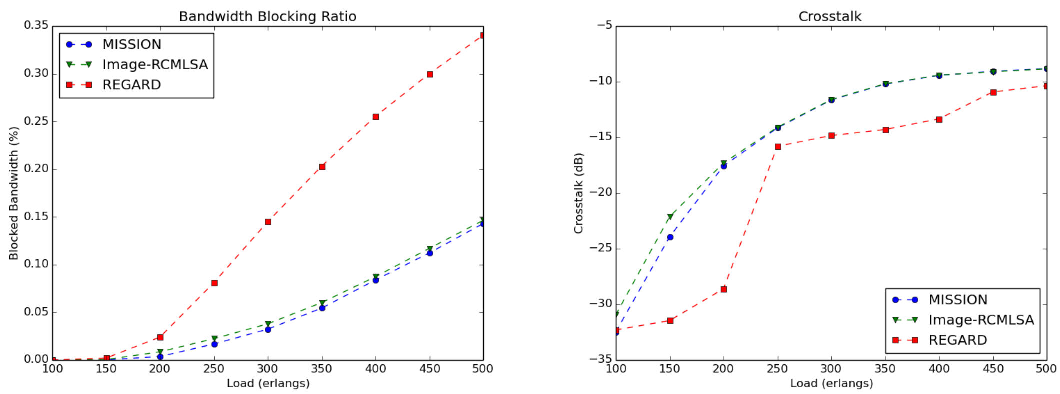


Figure 5. NSF topology

4.2. Results

We compared the aforementioned metrics in three scenarios and compiled the results in Figures 6(a), 6(b), 7(a), and 7(b). The curve labeled as MISSION represents the performance of the proposed approach in this paper, and the remaining curves represent two algorithms from the literature, the Image-RCMLSA [Moura and da Fonseca 2018] and the REGARD[Rodrigues et al. 2021] algorithms, as labeled in the graphs.

The algorithm in [Moura and da Fonseca 2018] was the source of the training data employed in the consolidation of the proposed model, thus working with the same networking features and providing the most competitive baseline for comparison. The approach proposed by [Rodrigues et al. 2021] uses an entirely different heuristic for resource allocation while employing the very same metrics used in our approach for resource allocation optimization.



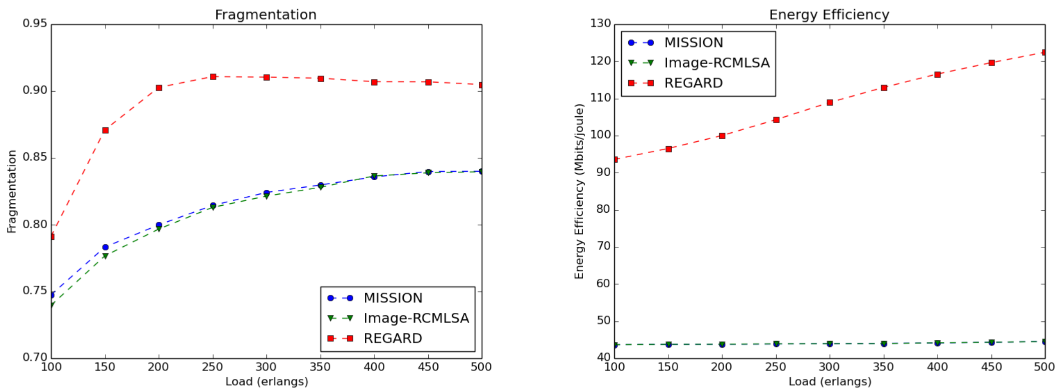
(a) Bandwidth Blocking Ratio (b) Crosstalk

Figure 6. Results

Figure 6(a) shows the bandwidth blocking ratio across all loads tested. It shows that our approach was consistently better at reducing blocking across all loads compared to the Image-RCMLSA algorithm, even if by a small margin, and it performed considerably better than the REGARD approach across all loads in a same network model.

Figure 6(b) shows the crosstalk across all loads tested. Regarding this metric, our approach was shown to perform better than the Image-RCMLSA at loads of 100 up to 250 erlangs, and performed competitively at loads of 250 up to 500 erlangs, never performing worse than the benchmarking algorithm. The REGARD algorithm showed better crosstalk performance across all loads, this is due to the REGARD algorithm generating more BBR and consequently producing less resources on the network and consequently less XT.

Figure 7 shows the fragmentation ratio and energy efficiency. Figure 7(a) shows the spectrum fragmentation across all loads tested. Our approach managed to surpass the Image-RCMLSA at loads up to about 250 erlangs and remained at competitive levels at all loads up to 500. It also performed significantly better than the REGARD algorithm at all loads.



(a) Fragmentation Ratio.

(b) Energy Efficiency

Figure 7. Results

Figure 7(b) shows the energy efficiency of the approaches compared. Given that our approach does not expressly rely on distance as its main component for path selection, but rather on a set of metrics and network features, it could be argued that bandwidth blocking, fragmentation, and crosstalk reduction would be achieved at the cost of a certain amount of deterioration to the energetic efficiency of the network.

5. Conclusions

In this paper we proposed the MISSION approach for traffic-aware link routing in SDM-EON networks. This approach aims at reducing request blocks while maintaining acceptable levels of crosstalk and fragmentation. This approach was compared to a competing heuristic approach in the literature and showed competitive results, which demonstrate that ML-based traffic-aware approaches to the RMSCA problem can be a highly adaptive, data-based, alternative approach to the current static heuristic approaches.

The next development in this research should involve replacing the distance, used as weight in the vertices that connect the links in the network, with a more meaningful metric, i.e. a metric calculated by means of an ML model trained with the network data collected for this study. By replacing the distance in the allocation process, the Dijkstra algorithm could estimate the optimal path based on a much more reliable metric than only distance, which could lead to potencily greater reductions in MBBR, XT and fragmentation.

Aknowledgments

This work was supported by grant #2020/05054-5, São Paulo Research Foundation (FAPESP), and by the Institutional Program of Scientific initiation grant.

References

Amirabadi, M. (2019). A survey on machine learning for optical communication [machine learning view]. *arXiv preprint arXiv:1909.05148*.

Beyragh, A. A., Rahbar, A. G., Ghazvini, S.-M. H., and Nickray, M. (2019). If-rsca: intelligent fragmentation-aware method for routing, spectrum and core assignment in space division multiplexing elastic optical networks (sdm-eon). *Optical Fiber Technology*, 50:284–301.

Cisco, U. (2018). Cisco visual networking index (vni). complete forecast update, 2017-2022. Accessed: 2020-07-03.

El Naqa, I. and Murphy, M. J. (2015). What is machine learning? In *machine learning in radiation oncology*, pages 3–11. Springer.

Grinberg, M. (2018). *Flask web development: developing web applications with python*. " O'Reilly Media, Inc.".

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825):357–362.

Khan, F. N., Fan, Q., Lu, J., Zhou, G., Lu, C., and Lau, P. T. (2020). Applications of machine learning in optical communications and networks. In *Optical Fiber Communication Conference*, pages M1G–5. Optical Society of America.

Moura, P. M. and da Fonseca, N. L. S. (2018). Routing, core, modulation level, and spectrum assignment based on image processing algorithms. *J. Opt. Commun. Netw.*, 10(12):947–958.

Moura, P. M. and Drummond, A. C. FlexGridSim: Flexible Grid Optical Network Simulator. http://www.lrc.ic.unicamp.br/FlexGridSim/.

Oliveira, H. M. N. S. and Fonseca, N. L. S. (2019). Multipath routing, spectrum and core allocation in protected sdm elastic optical networks. In *2019 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE.

Paira, S., Halder, J., Chatterjee, M., and Bhattacharya, U. (2020). On energy efficient survivable multipath based approaches in space division multiplexing elastic optical network: Crosstalk-aware and fragmentation-aware. *IEEE Access*, 8:47344–47356.

pandas development team, T. (2020). pandas-dev/pandas: Pandas.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Rodrigues, E., Figueiredo, G., Santi, J., and Oliveira, H. (2021). Roteamento e alocação de núcleo e espectro com ciência de fragmentação e crosstalk em sdm-eon. In *Anais do XXXIX Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 546–559, Porto Alegre, RS, Brasil. SBC.

Trindade, S. and da Fonseca, N. L. (2021). Machine learning for spectrum defragmentation in space-division multiplexing elastic optical networks. *IEEE Network*, 35(1):326–332.

Trindade, S. and da Fonseca, N. L. S. (2019). Proactive fragmentation-aware routing, modulation format, core, and spectrum allocation in eon-sdm. In *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, pages 1–6.

Van Rossum, G. (2020). *The Python Library Reference, release 3.8.2*. Python Software Foundation.

Yang, S., Yang, L., Luo, F., Wang, X., Li, B., Du, Y., and Liu, D. (2021). Multi-channel multi-task optical performance monitoring based multi-input multi-output deep learning and transfer learning for sdm. *Optics Communications*, 495:127110.

Yao, Q., Yang, H., Yu, B., Yu, A., Bai, W., and Zhang, J. (2018a). Spectrum optimization for resource reservation based on transductive transfer learning in space division multiplexing elastic optical networks. In *2018 European Conference on Optical Communication (ECOC)*, pages 1–3.

Yao, Q., Yang, H., Zhu, R., Yu, A., Bai, W., Tan, Y., Zhang, J., and Xiao, H. (2018b). Core, mode, and spectrum assignment based on machine learning in space division multiplexing elastic optical networks. *IEEE Access*, 6:15898–15907.

Yousefi, F., Ghaffarpour Rahbar, A., and Ghadesi, A. (2020). Fragmentation and time aware algorithms in spectrum and spatial assignment for space division multiplexed elastic optical networks (sdm-eon). *Computer Networks*, 174:107232.

Yousefi, F. and Rahbar, A. G. (2020). Novel crosstalk, fragmentation-aware algorithms in space division multiplexed-elastic optical networks (sdm-eon) with considering physical layer security. *Optical Switching and Networking*, 37:100566.

Zhao, Y., Hu, L., Zhu, R., Yu, X., Wang, X., and Zhang, J. (2018). Crosstalk-aware spectrum defragmentation based on spectrum compactness in space division multiplexing enabled elastic optical networks with multicore fiber. *IEEE Access*, 6:15346–15355.

Zhu, R., Samuel, A., Wang, P., Li, S., Oun, B. K., Li, L., Lv, P., Xu, M., and Yu, S. (2021). Protected resource allocation in space division multiplexing-elastic optical networks with fluctuating traffic. *Journal of Network and Computer Applications*, 174:102887.

Ítalo Brasileiro, Costa, L., and Drummond, A. (2020). A survey on challenges of spatial division multiplexing enabled elastic optical networks. *Optical Switching and Networking*, 38:100584.