On Global Fragmentation Metrics as Proxy for Network Blocking: Correlation, Detection and Prediction

Hakim Mellah
electrical engineering dept.
Ecole Polytehnique, Montreal
hakim.mellah@polymtl.ca

Étienne Payette-Toupin
electrical engineering dept.
Ecole Polytehnique, Montreal
etienne.payette-toupin@polymtl.ca

Renaud Lespérance
electrical engineering dept.
Ecole Polytehnique, Montreal
renaud.lesperance@polymtl.ca

Scott Kohlert
Ciena
Ciena, Ottawa
skohlert@ciena.com

Brunilde Sansò
electrical engineering dept.
Ecole Polytehnique, Montreal
brunilde.sanso@polymtl.ca

***Abstract*—Elastic Optical Networks (EONs) are challenged by spectrum fragmentation, which can obstruct the establishment of new connections. While the concept of fragmentation gaps is relatively simple, determining an accurate metric to quantify fragmentation remains complex, and numerous metrics have been proposed to address this.**

In this work, we analyze the effectiveness of various fragmentation metrics in understanding and forecasting network blocking. We begin by evaluating the correlation between common global fragmentation metrics and blocking probability, finding a relationship between the two. Through the application of a Random Forest model, we identify the most useful metrics for anticipating blocking caused by fragmentation. Our classification model further assesses whether these metrics can accurately detect current blocking and predict future blocking events. The findings reveal that although global fragmentation metrics perform better than chance, they are not yet highly reliable predictors, opening avenues for future exploration of more detailed fragmentation metrics.

***Index Terms*—Elastic Optical Networks (EON), Bandwidth Fragmentation, Fragmentation Metrics, Blocking Rate.**

I. Introduction

Forthcoming high bandwidth and stringent applications increasingly need the flexibility of Elastic Optical Networks (EONs) [1], [2], [3]. In EONs the optical spectrum is divided in frequency slots (FSs) [4] and, depending on the modems, different number of contiguous slots can be allocated to an incoming connection. When connections terminate, slots become available and spectrum holes are created. Because the allocated frequency slots for future connections must be contiguous (on each link) and continuous (along the entire path), the holes left by departing connections may not be sufficient for incoming connections, creating fragmentation and increasing the probability of network blocking [5].

While the seriousness of fragmentation can be visually capted, it is not easy to quantify. As a result, several metrics have been proposed such as External Fragmentation (EF), Shannon Entropy (SE), Access Blocking Probability (ABP), Root Mean Square Factor (RMSF) and Wasted Slots (WS) [3]. WS is a horizontal metric, which can be used to measure the spectrum continuity, while the others are vertical metrics that can be used to measure the spectrum contiguity. The literature also proposes a combination of two metrics such as RMSF and WS [3].

We intuitively know that a fragmented spectrum produces blocking, therefore implementing defragmentation algorithms is an important step to reduce blocking. The question that remains open, however, is what type of metric should be used to steer those defragmentation algorithms? Ideally, because blocking is the phenomenon that we want to avoid, the blocking probability should be the metric to be used. Unfortunately, in the context of EON, no analytic formulas exist to exactly calculate the probability of blocking connections based on the current capacity and on-going demand. Thus, measuring blocking can only be done *a-posteriori* after it is actually installed in the network. Fragmentation metrics, on the other hand, are known *a-priori* and can therefore be used as a monitoring mechanism before important blocking takes place. Therefore an implicit agreement in the literature has been that fragmentation metrics can be used as proxy for network blocking.

In recent years, there has been an increased interest in the problems surrounding fragmentation and defragmentation for EON [6] and proposals for different fragmentation metrics [7] to be able to quantify fragmentation for different types of traffic and situations. There has also been a surge in studies of Artificial Intelligence methods applied to Optical Networks and on how to implement those methods to be able to determine the best time to trigger network defragmentation [8], [9].

However, to the best of our knowledge, there has not been a formal correlation analysis and feature selection to investigate:

- if fragmentation metrics can really be used as a proxy for network blocking;
- if so, which of the most popular metrics would be the most appropriate;

- how to incorporate it for blocking detection and prediction

The objective of this paper is to shed light on the above-mentioned issues. Thus, the original contributions of this paper are:

- A comprehensive correlation analysis between global network blocking and single and composed fragmentation metrics evaluated at the network level.
- A feature importance analysis based on Random Forest to assess the best fragmentation metrics to be used as network blocking proxies
- A classification analysis to detect and predict a network blocking situation.

II. Methodology

To answer to the three objectives that we presented at the end of Section I, we first created an EON network simulator that would provide blocking as well as fragmentation à posteriori data. Next a correlation analysis was performed for different types of single or compound fragmentation metrics for the network as a whole with global blocking. The next step was a feature importance analysis based, followed by the use of Machine Learning Models for Blocking Detection and Blocking Prediction. We now provide some details for this general methodology.

A. Fragmentation metrics

In this paper we will consider that all network nodes are purely optical, then fragmentation can be caused due to two fundamental constraints which must always be respected in this type of optical network.

- **Contiguity:** A connection must use a block of contiguous set of frequency slots.
- **Continuity:** A connection on an optical channel must keep the same wavelengths (or FSs) through its entire path from origin to destination.

In this work, we consider the fragmentation metrics:

1 External Fragment (EF)
2 Shannon's Entropy (SE)
3 Access Blocking Probability (ABP)
4 Root Square Mean Factor (RSMF)
5 Wasted Slots (WS)

In addition to the previously mentioned simple fragmentation metrics, we have combined more than one metric in order; for instance; to quantify vertical and horizontal fragmentation in the same metric. For example, the combination of WS (F_{WS}) and RMSF (F_{RMSF}) metrics is obtained by the multiplication:

$$F_{RMSF-WS} = F_{RMSF} \times F_{WS} \qquad (1)$$

B. Simulation

We created a thorough network simulator that accepts any optical network topology and restrictions. In the simulator, connection demands arrive randomly at any node and can be destined to any node in the network. They are terminated after the end of their service time, that is also a random variable. Both fragmentation metrics and blocking rates are computed during the simulation.

C. Correlation and feature importance

The study on correlation is carried out by computing the Pearson correlation coefficients for blocking versus the individual or composed fragmentation metric for the overall network.

After the correlation study, we plan to carry out a feature importance analysis for the different fragmentation metrics to be used in the subsequent blocking detection and prediction steps.

D. Blocking Detection and Prediction

These tests will determine if a global value of fragmentation can indeed be used as a proxy for detection and prediction of blocking in the network. For this, classification problems will be solved where the two classifying categories are "Blocking" or "Not Blocking". In other words, we are not concerned with the level of blocking that has been experimented but just if there is any network blocking being installed in the system.

For detection, we will train the model considering several number of features and will evaluate different performance measures. For prediction, we will train the model with a vector of statistics from a window of instants of time.

As a performance measures of the ML models, we have used the following:

1 Accuracy,
2 Precision,
3 Recall,
4 f1_score.

III. Results and Discussions

A. Network topology and demands

In this work, two realistic networks were used in the simulation: the North American (NA) optical network provided by our industrial collaborator Ciena shown in Figure 1 and the Germany network [10]. The nodes in NA aare gridless CDC (colorless, directionless, contentionless) or CD (colorless, directionless) ROADMs (Reconfigurable Optical Add/Drop Multiplexer), or DGE (Dynamic Gain Equalizer) sites. Each link has a bandwidth of $4800\ GHz$, subdivided into 768 FS, of $6.25\ GHz$ each. Furthermore, each link may consist of several segments and each segment has a number of amplifiers.

In addition to that, three types of modems corresponding to different applications and having different bandwidth requirements are considered. These modems are the WaveLogic 5 Extreme (WL5e) for long-haul 400GbE client transport, the WaveLogic Ai (WLAi) for hard-to-predict bandwidth demands, and the WaveLogic 3 (WL3) to meet the need of 10GbE to 100GbE clients. To meet the bandwidth of these different applications, the WL5e, WLAi and WL3 require bandwidths of 118.75 GHz, 75 GHz and 50 GHz respectively.

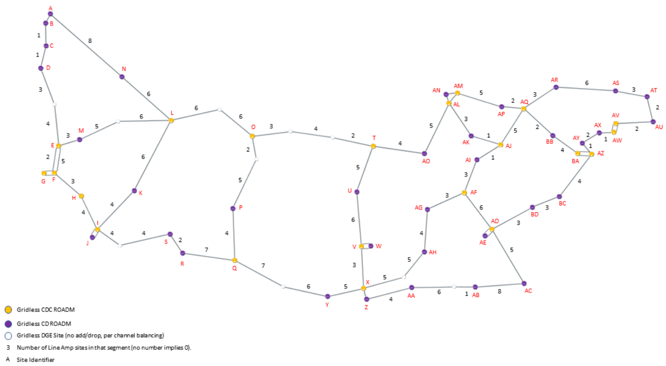


Fig. 1: North American optical network example provided by Ciena

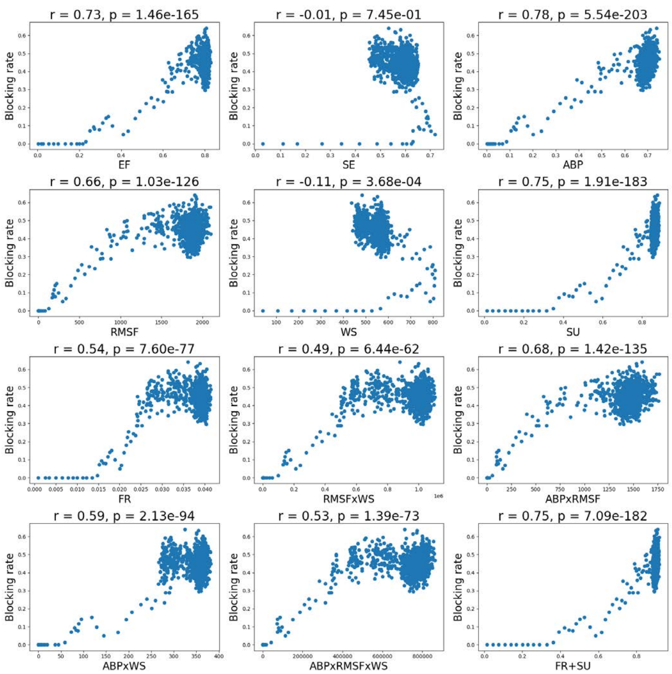


Fig. 2: Correlation between blocking rate and fragmentation metrics in use case 2 (NA with medium traffic intensity)

Furthermore, First Fit (FF) was used for routing and spectrum allocation (RSA) [6].

Moreover, two intensities of load are considered in each network: medium and high loads.

B. Correlation Tests

We started our work by studying the correlation between the blocking rate and the different fragmentation metrics.

Figs.2-5 show the distribution of the blocking with respect to each fragmentation metric for the networks used in this study. The correlation coefficients (the r and p values from the Pearson's correlation tests) are shown on top of each figure.

The results from the figures first show that the correlation coefficients depend on the type of metric. Moreover, the r and p values suggest that even though the two quantities (network blocking and global fragmentation) are not perfectly linearly correlated they are still directly related.

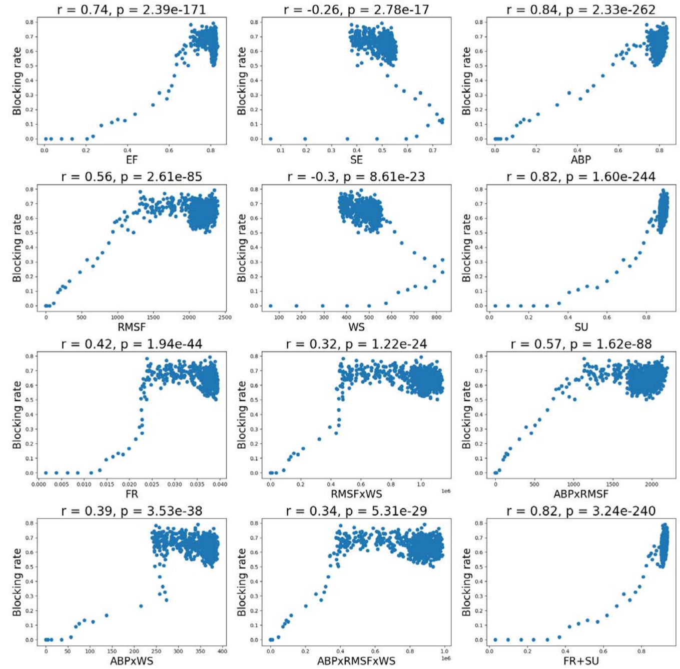


Fig. 3: Correlation between blocking rate and fragmentation metrics in use case 3 (NA with high traffic intensity)

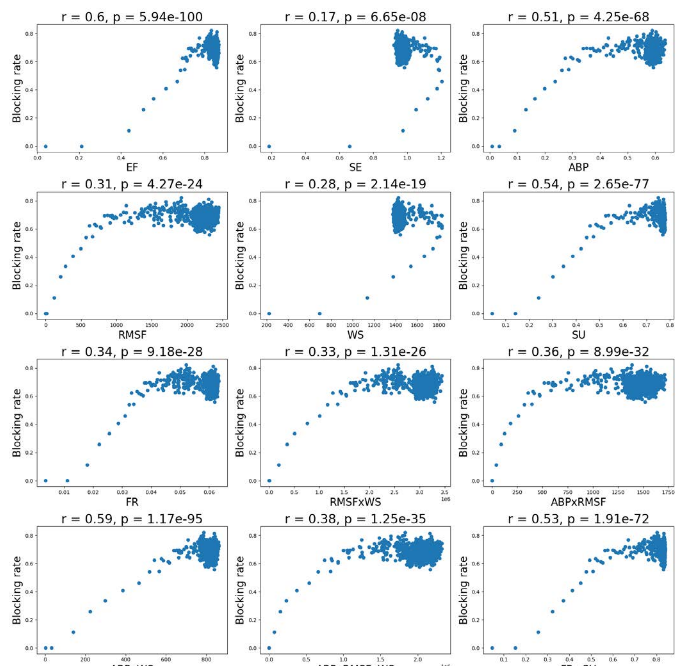


Fig. 4: Correlation between blocking rate and fragmentation metrics in use case 8 (Germany with medium traffic intensity)

C. Metric Feature Importance

To choose the best fragmentation metric(s) to use for the prediction of the blocking, we used a Random Forest (RF)

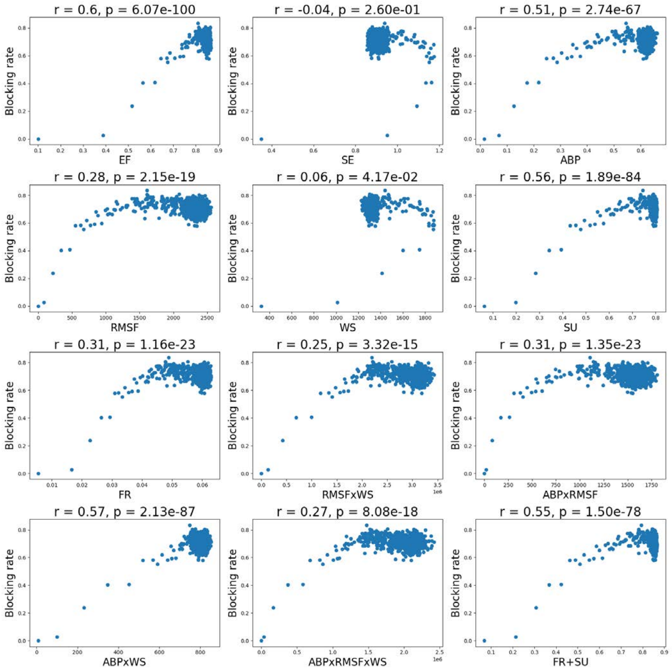
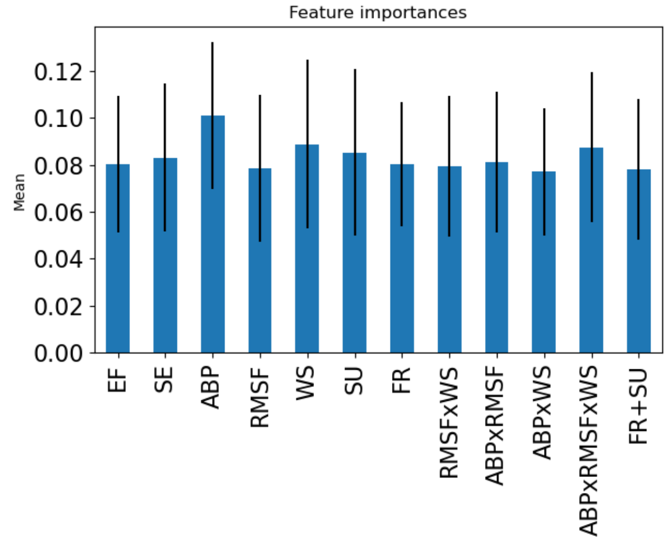
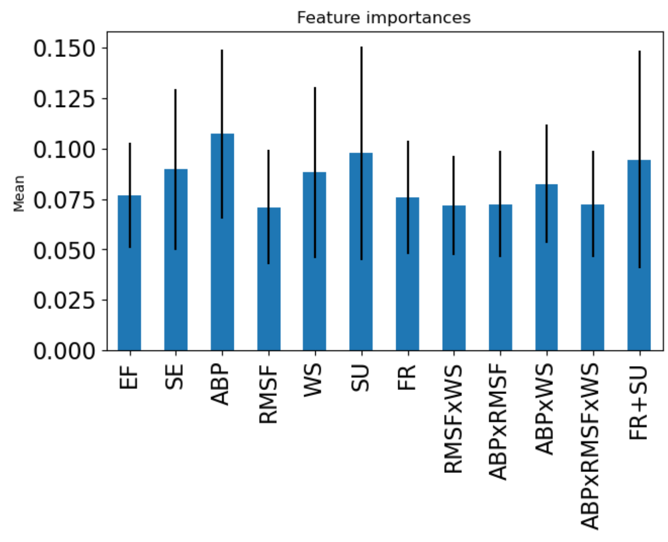


Fig. 5: Correlation between blocking rate and fragmentation metrics in use case 9 (Germany with high traffic intensity)



(a) Medium traffic load



(b) High traffic load

Fig. 6: Feature importance in NA network

model to calculate the average of the feature importance of each fragmentation metric. The results of these tests are shown in Figs.6-7. The figures show that all features have similar importance and none of them dominates the others. This suggests that the features are redundant and using one or few of them would lead to the same results as using all the features.

To further investigate on the feature importance, Table I presents the performance results of the blocking detection using 1, 2, 3 and all features for each case.

In our study, we use a blocking threshold to consider he network being in either a blocking situation, when the average blocking in the network is above the threshold or in a non-blocking situation otherwise. This results in considering the problem as a classification problem. The blocking threshold is chosen for each case to have both classes (blocking and non blocking) to have similar occurrences.

The results show that the difference between using 1 feature and all 12 features vary from 8% (precision in Germany medium) and 26% (recall in NA high).

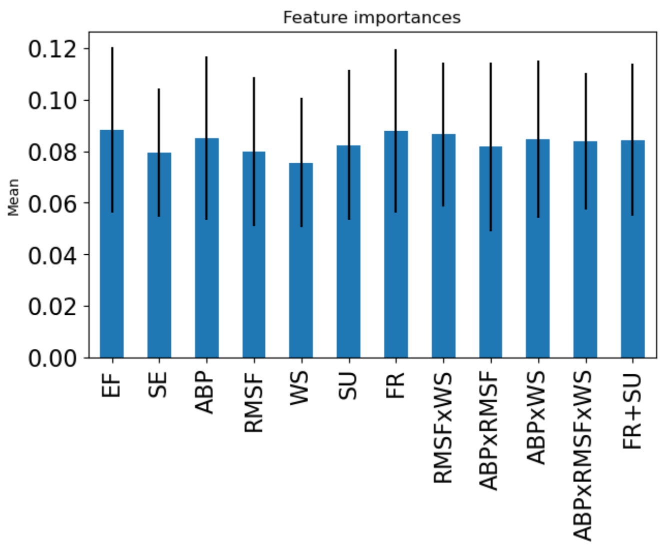
D. Blocking Prediction using Statistical Vector

In order to help the model make better predictions, more data about the features is to be provided. In our case, we ran simulations and generated global fragmentation metrics. Of course, it could also be possible to generate fragmentation metrics per link and per path, but that would denature the objective of our study that is to inquire if global metrics are enough to predict blocking.

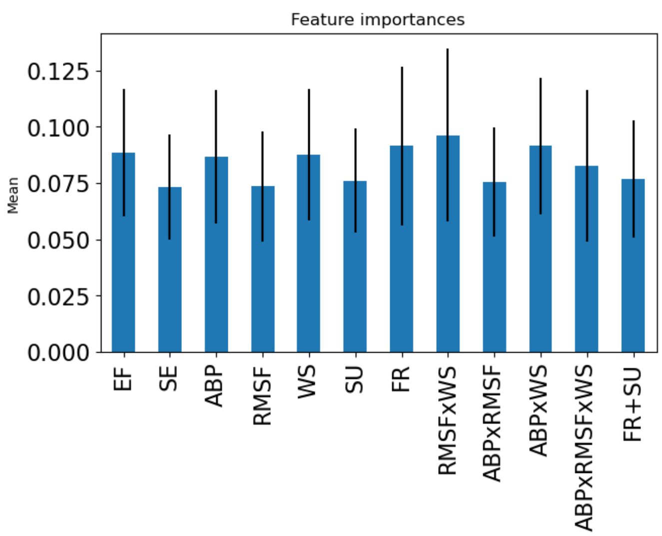
We propose to use the general fragmentation metrics and produce statistical vectors at each time step by considering a window of previous metrics. For example, if we use a window of size w, then at time step t_k, we use the metrics obtained at time steps $t_k, t_{k-1}, ..., t_{k-w}$, $k > w$ and then generate a vector of statistics of these measure, such as the max, the min, the mean, the first derivatives ($\frac{f_{k-w+1}-f_{k-w}}{dt}, ..., \frac{f_k - f_{k-1}}{dt}$, where f_k is the fragmentation metric at time step k and $dt = t_k - t_{k-1}$).

Table II illustrates the detection performance results as a function of the window size (w). In this table, we compare the performance using one feature with and without statistical vector. We have used four widow sizes (2, 3, 4, and 5). Window size 1 correspond to using the metric without any statistics. As a reference, we added the case of using all features to each use case.

As a general comment, using a vector of statistics improves the detection performance. There are some few exception such



(a) Medium traffic load



(b) High traffic load

Fig. 7: Feature importance in Germany network

Network	Features	Performance			
		Accuracy	Precision	Recall	F1
NA, medium	All	0.59	0.58	0.6	0.59
	3	0.59	0.58	0.63	0.6
	2	0.63	0.65	0.59	0.62
	1	0.49	0.49	0.52	0.5
NA, High	All	0.67	0.66	0.73	0.7
	3	0.64	0.65	0.68	0.66
	2	0.65	0.65	0.71	0.67
	1	0.58	0.61	0.54	0.57
Germany, medium	All	0.63	0.63	0.72	0.67
	3	0.65	0.66	0.68	0.67
	2	0.61	0.63	0.66	0.64
	1	0.55	0.58	0.58	0.58
Germany, High	All	0.62	0.7	0.6	0.65
	3	0.63	0.72	0.61	0.66
	2	0.65	0.73	0.62	0.67
	1	0.52	0.61	0.47	0.53

TABLE I: The performance results of the cases used in the study

Network	Windows	Performance			
		Accuracy	Precision	Recall	F1
NA, medium	All features	0.59	0.58	0.6	0.59
	Window = 1	0.49	0.49	0.52	0.5
	Window = 2	0.53	0.5	0.48	0.49
	Window = 3	0.55	0.51	0.57	0.53
	Window = 4	0.59	0.55	0.66	0.6
	Window = 5	0.61	0.62	0.65	0.63
NA, high	All features	0.67	0.66	0.73	0.7
	Window = 1	0.58	0.61	0.54	0.57
	Window = 2	0.59	0.61	0.62	0.61
	Window = 3	0.6	0.59	0.62	0.61
	Window = 4	0.6	0.53	0.73	0.62
	Window = 5	0.58	0.6	0.64	0.62
Germany, medium	All features	0.63	0.63	0.72	0.67
	Window = 1	0.55	0.58	0.58	0.58
	Window = 2	0.51	0.6	0.53	0.56
	Window = 3	0.58	0.64	0.6	0.62
	Window = 4	0.52	0.62	0.53	0.57
	Window = 5	0.52	0.62	0.53	0.57
Germany, high	All features	0.62	0.7	0.6	0.65
	Window = 1	0.52	0.61	0.47	0.53
	Window = 2	0.59	0.63	0.59	0.61
	Window = 3	0.61	0.64	0.62	0.63
	Window = 4	0.63	0.68	0.56	0.61
	Window = 5	0.63	0.68	0.56	0.61

TABLE II: The performance results using a statistical vector

as the recall in NA medium case, where using a vector of statistics degrades the performance of the recall.

It can also be seen from the table that this improvement achieved saturation after $w = 5$ in most cases.

E. ML models for Blocking Future Prediction

In this section, we present the results obtained from ML models to predict future blocking in the network.

We have trained the model to predict the blocking at some forward steps in the future. For instance, if the step size is s, we train the model to predict the blocking at time step t_{k+s} given the measurement of a single fragmentation metric at time step t_k.

Table III and Table IV show the prediction performance results as a function of the future step size. The first tables is for the near future (up to 10 time steps) and the second table is for the far future (up to 100 time steps).

The results suggest that the prediction capabilities of the model from global fragmentation metrics is better than pure chance but is far from being a strong predictor. Nevertheless they open the path for more refined fragmentation metrics to be used in global blocking prediction.

IV. Conclusion

While the EON literature has often associated fragmentation with network blocking, this paper took a step forward by exploring the strength of this relationship. Through extensive simulations, we first examined the correlation between various global fragmentation metrics and blocking rates, confirming a clear relationship between the two. This was followed by

Network	Steps	Performance			
		Accuracy	Precision	Recall	F1
NA, medium	0	0.49	0.49	0.52	0.5
	2	0.55	0.53	0.51	0.52
	4	0.5	0.46	0.47	0.47
	6	0.57	0.6	0.56	0.58
	8	0.53	0.52	0.51	0.51
	10	0.56	0.53	0.6	0.56
NA, medium	0	0.58	0.61	0.54	0.57
	2	0.59	0.61	0.56	0.59
	4	0.5	0.44	0.52	0.48
	6	0.52	0.53	0.54	0.53
	8	0.63	0.74	0.62	0.67
	10	0.46	0.53	0.47	0.5
Germany, medium	0	0.52	0.54	0.59	0.57
	2	0.63	0.71	0.62	0.66
	4	0.52	0.6	0.59	0.6
	6	0.52	0.58	0.53	0.56
	8	0.54	0.61	0.56	0.58
	10	0.5	0.59	0.46	0.52
Germany, high	0	0.52	0.61	0.47	0.53
	2	0.61	0.65	0.59	0.62
	4	0.54	0.56	0.56	0.56
	6	0.53	0.62	0.41	0.49
	8	0.5	0.56	0.47	0.51
	10	0.52	0.61	0.47	0.53

TABLE III: The performance for future prediction to 10 steps

Network	Steps	Performance			
		Accuracy	Precision	Recall	f1
NA, medium	20	0.53	0.6	0.49	0.54
	40	0.51	0.55	0.56	0.56
	60	0.55	0.55	0.64	0.59
	80	0.42	0.42	0.47	0.44
	100	0.54	0.55	0.49	0.52
NA, high	20	0.51	0.55	0.49	0.52
	40	0.51	0.47	0.58	0.52
	60	0.53	0.64	0.51	0.57
	80	0.57	0.6	0.64	0.62
	100	0.5	0.47	0.57	0.51
Germany, medium	20	0.46	0.46	0.53	0.49
	40	0.49	0.5	0.61	0.55
	60	0.55	0.59	0.56	0.57
	80	0.46	0.52	0.49	0.5
	100	0.51	0.49	0.49	0.49
Germany, high	20	0.52	0.53	0.53	0.53
	40	0.53	0.59	0.53	0.56
	60	0.6	0.62	0.56	0.59
	80	0.46	0.42	0.52	0.47
	100	0.51	0.49	0.55	0.52

TABLE IV: The performance for future prediction to 100 steps

the development of machine learning models to assess global fragmentation predicting power.

Our machine learning models revealed that while global fragmentation metrics provide useful information, they may not be the strongest predictors for network blocking. Building on these findings, we are actively exploring whether more granular measurements, such as per-link or per-path fragmentation, could offer greater predictive power for accurately forecasting network blocking.

REFERENCES

[1] A. S. Santos, J. de Santi, G. B. Figueiredo, and B. Mukherjee, "Application-aware service degradation in elastic optical networks," *IEEE Transactions on Network and Service Management*, vol. 19, no. 2, pp. 949–961, 2022.

[2] W. Halabi, "Traffic aggregation in ip-based wdm vs. ip-based elastic optical networks," in *2020 International Conference on Promising Electronic Technologies (ICPET)*, 2020, pp. 140–143.

[3] P. Lechowicz, "Regression-based fragmentation metric and fragmentation-aware algorithm in spectrally-spatially flexible optical networks," *Computer Communications*, vol. 175, pp. 156–176, Jul. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140366421001961

[4] S. Fernández-Martínez, B. Barán, and D. P. Pinto-Roa, "Spectrum defragmentation algorithms in elastic optical networks," *Optical Switching and Networking*, vol. 34, pp. 10–22, 2019.

[5] P. M. Moura, N. L. S. d. Fonseca, and R. A. Scaraficci, "Fragmentation aware routing and spectrum assignment algorithm," in *2014 IEEE International Conference on Communications (ICC)*, 2014, pp. 1137–1142.

[6] B. C. Chatterjee, W. Fadini, and E. Oki, "A spectrum allocation scheme based on first-last-exact fit policy for elastic optical networks," *J. Netw. Comput. Appl.*, 2016.

[7] P. Lechowicz, M. Tornatore, A. Wlodarczyk, and K. Walkowiak, "Fragmentation metrics and fragmentation-aware algorithm for spectrally/spatially flexible optical networks," *Journal of Optical Communications and Networking*, vol. 12, no. 5, pp. 133–145, 2020.

[8] R. Gao, L. Liu, X. Liu, H. Lun, L. Yi, W. Hu, and Q. Zhuge, "An overview of ml-based applications for next generation optical networks," *Science China Information Sciences*, vol. 63, pp. 1–16, 2020.

[9] E. Dávalos, J.-L. Enciso, N. Silva, J. Pinto-Ríos, and A. Leiva, "Triggering strategy for defragmentation process in elastic optical networks using machine learning techniques," *ICT Express*, vol. 9, no. 5, pp. 890–895, 2023.

[10] sndlib. Germany Network. [Online]. Available: https://sndlib.put.poznan.pl/home.action