Dynamic Power Pre-adjustments with Machine Learning that Mitigate EDFA Excursions during Defragmentation

Yishen Huang, Patricia B. Cho, Payman Samadi, Keren Bergman
Department of Electrical Engineering, Columbia University, New York, NY,10027, USA
y.huang@columbia.edu

Abstract: We examine EDFA power excursions during three defragmentation methods of flexgrid super-channels. Using a machine learning approach, we demonstrate automatic and dynamic adjustments of pre-EDFA power levels, and show the mitigation of post-EDFA power discrepancy among channels by over 62%.
OCIS codes: (060.4510) Optical communications; (060.4256) Networks, network optimization

1. Introduction

In flexgrid networks implementing super-channels, dynamic traffic arrivals and departures yield a fragmented spectrum with increased network blocking probability [1]. Hence, spectral defragmentation is performed to adjust channel wavelength assignment and re-groom contiguous bandwidths. While dynamic defragmentation has been explored [2–4], previous studies focus on enabling technologies and network layer considerations without addressing the concern of power dynamics associated with dynamic networking [5]. In particular, as a result of the non-flat gain-tilt and automatic gain control mechanisms [6], the power excursion problem in Erbium Doped Fiber Amplifiers (EDFA) remains unsolved for dynamically changing spectral configurations. In [7] we examined the impact on EDFA power excursions when fixgrid channels are added or dropped. Due to the channel dependence of EDFA power excursions, changing a super-channel's location interferes with the network's power dynamics. Since super-channels consist of multiple contiguous sub-channels, they may induce greater power changes in comparison to individual fixgrid channels and present a greater challenge to maintaining post-EDFA power stability.

The defragmentation methods can be categorized into three classes: Hop [2], Make-before-Break (MbB) [3], and Sweep [4]. In Hop, a new channel is established while dropping the original channel simultaneously. In MbB, a new channel is created, allowing the traffic to be re-established before dropping the original channel. In Sweep, the channel wavelength is shifted gradually at the spectral granularity of the equipment without disruption to the traffic. Each of the three methods induces a unique power impact on the EDFAs. In this work, we characterize the channel-dependence of power excursions for all three methods with Machine Learning (ML), and introduce an adaptive strategy for automatic power adjustments during the defragmentation process. We experimentally demonstrate the mitigation of post-EDFA power discrepancy during and after the defragmentation process.

2. Approach

We define *undesired* power excursions as increases in the discrepancy of post-EDFA channel powers. Adjustments on pre-EDFA channel powers has been shown to effectively compensate for the post-EDFA discrepancy [8]. In order to make adjustments adaptive to the power dynamics effects specific to a channel's location, two metrics are needed: magnitude and correlation. The magnitude specifies the extent to which a channel's location contributes to the the post-EDFA power discrepancy. The correlation specifies whether a rise in the channel's pre-EDFA power would increase or decrease the discrepancy among post-EDFA powers. We demonstrate that both metrics can be characterized for cascaded EDFAs with low-complexity ML regression and classification models, which then allow one-step adjustments in real time to maintain power stability throughout the defragmentation process.

2.1. Magnitude of impact

We formulate a regression problem to determine the contribution of each spectral granularity to the post-EDFA discrepancy. Using a Ridge Regression (RR) model, we determine a set of linear weights w_{RR} associated with the sub-channels. Fig. 1a shows the weights learned by RR for the experimental 24-channel Wavelength-division Multiplexing (WDM) system, shown in Fig. 1b. A positive weight indicates that the channel would increase the post-EDFA discrepancy when turned ON, and a negative weight indicates a decrease. The magnitude of each weight shows the relative extent of the sub-channel's contribution. All channels shown in Fig. 1a have the same pre-EDFA power, and the discrepancy is solely due to the power excursions induced by the cascaded EDFAs.

2.2. Correlation of impact

Studies in [6] show that channels with both high and low post-EDFA powers can result in undesired power excursions in cascaded EDFAs. A rise in a channel's pre-EDFA power may increase (correlation = +1) or decrease (correlation

= -1) the post-EDFA discrepancy. Hence, a channel's correlation determines the appropriate bias on its pre-EDFA power. We formulate a Logistic Regression (LR) model to predict a channel's correlation given the current spectral usage. The distribution $P(s_x|\{x^-\})$ is learned, where s_x is the correlation specific to channel x and the set $\{x^-\}$ indicates the ON/OFF states of all other channels in the same light path. The training process learns a set of weights, w_{LR}^x, for every channel x, which is used in a sigmoid function to determine the correlation.

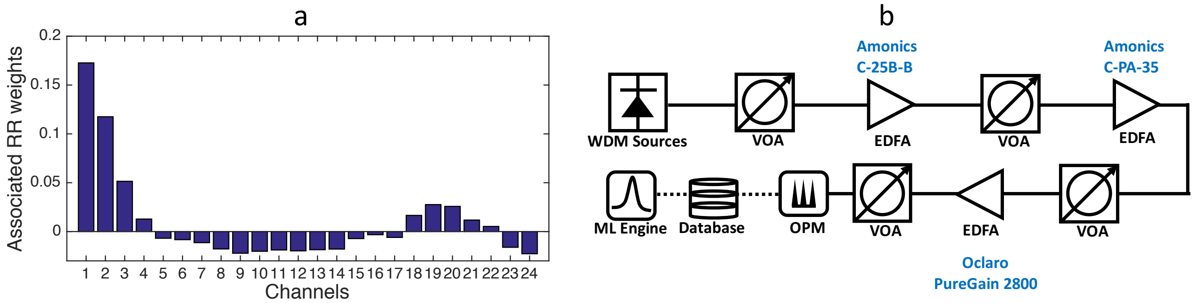

Fig. 1. (a) Weights associated with each WDM sub-channel from the trained RR model. (b) Component diagram of the experimental setup

3. Algorithm and experimental validation

We collected 800 training scenarios and 200 testing scenarios with randomly initiated ON and OFF channels. The ON/OFF states of channels and their post-EDFA powers were recorded. Once trained, RR achieves a mean square error (MSE) of 0.004 when predicting the post-EDFA variance of the testing set; and LR achieves a 98% classification accuracy against the testing set. Validated by the high accuracy in their predictions, the ML models are incorporated into Alg. 1-3 to determine the adaptive pre-EDFA power adjustments for the newly established super-channel.

Alg. 1. Hop Defragmentation Optimization

IF RR, LR are untrained:
  Train RR, LR
  Save w_{RR}, $w_{LR}^{\{x\}}$
End IF
SET ch_{sup}^{new} # the new super-channel location in sub-channel indices
SET $\{x\}^{new}$ # the new spectral configuration
FOR each ch in ch_{sup}^{new}:
  $magnitude^{ch} = w_{RR}(ch)$
  $correlation^{ch} = sign\left(sigmoid\left(w_{LR}^{ch} \cdot \{x\}^{new}\right) - 0.5\right)$
  IF $magnitude^{ch} > 0$:
    $powerAdjust^{ch} = -correlation^{ch} \times \frac{magnitude^{ch}}{max(|w_{RR}|)} \times powerUnit$
  $power^{ch} \mathrel{+}= powerAdjust^{ch}$
End

Alg. 3. Sweep Defragmentation Optimization

IF RR, LR are untrained:
  Train RR, LR
  Save w_{RR}, $w_{LR}^{\{x\}}$
End IF
REPEAT until super-channel shifted to desired location:
  SET ch_{sup}^{next} # the next super-channel location after shifting by 1 step
  SET $\{x\}^{next}$ # the next spectral configuration
  FOR each ch in ch_{sup}^{next}:
    $magnitude^{ch} = w_{RR}(ch)$
    $correlation^{ch} = sign\left(sigmoid\left(w_{LR}^{ch} \cdot \{x\}^{next}\right) - 0.5\right)$
    IF $magnitude^{ch} > 0$:
      $powerAdjust^{ch} = -correlation \times \frac{magnitude^{ch}}{max(|w_{RR}|)} \times powerUnit$
    End IF
    $power^{ch} \mathrel{+}= powerAdjust^{ch}$
  END
END REPEAT

Alg. 2. MbB Defragmentation Optimization

IF RR, LR are untrained:
  Train RR, LR
  Save w_{RR}, $w_{LR}^{\{x\}}$
End IF
SET ch_{sup}^{new} # the new super-channel location in sub-channel indices
SET $\{x\}^{int}$ # the intermediate spectral configuration
SET $\{x\}^{fin}$ # the final spectral configuration
FOR each ch in ch_{sup}^{new}:
  $magnitude^{ch} = w_{RR}(ch)$
  $correlation^{ch} = sign\left(sigmoid\left(w_{LR}^{ch} \cdot \{x\}^{int}\right) - 0.5\right)$
  IF $magnitude^{ch} > 0$:
    $powerAdjust^{ch} = -correlation^{ch} \times \frac{magnitude^{ch}}{max(|w_{RR}|)} \times powerUnit$
  $power^{ch} \mathrel{+}= powerAdjust^{ch}$
DROP original super-channel
FOR each ch in ch_{sup}^{new}:
  $correlation^{ch} = sign\left(sigmoid\left(w_{LR}^{ch} \cdot \{x\}^{fin}\right) - 0.5\right)$
  IF $magnitude^{ch} > 0$:
    $powerAdjust^{ch} = -correlation^{ch} \times \frac{magnitude^{ch}}{max(|w_{RR}|)} \times powerUnit$
  $power^{ch} \mathrel{+}= powerAdjust^{ch}$
End

Leveraging the ML models, a single-step adjustment is computed based solely on the wavelength assignments of channels and without iterative tuning. The adjustments are localized to the shifted super-channel alone. We examine the feasibility of the algorithms with the cascaded EDFAs shown in Fig. 1b. The variable $powerUnit$ in the algorithms is set to 3dBm – the maximum tuning range of the channel launch power. Variable optical attenuators (VOA) are used to mimic 25dB loss in the fiber spans. Different models of EDFAs are employed to test the robustness of the algorithms against system complexity. A super-channel is simulated in the experiments by three spectrally contiguous channels. The WDM sources provide a total of 24 WDM channels from ITU-T grid 21-44, which we index as Ch. 1-24 respectively. In the following experiments, the super-channel is shifted as part of the defragmentation process beginning at Timestep 10.

Experiment 1 shifts a super-channel from location Ch. 9-11 to location Ch. 1-3. Ch. 13-24 are also ON, and all other channels are OFF. From Fig. 1a, this super-channel is shifted from a location that reduces power discrepancy to a location that significantly exacerbates it. Fig. 2 shows the difference in the post-EDFA variances between applying and not applying adaptive power. The variance of the post-EDFA power levels improves by 62.5% for Hop and MbB, effective immediately as the defragmentation begins. For Sweep, the longer transition period allows the adaptive power adjustments to achieve 75% improvement in power variance.

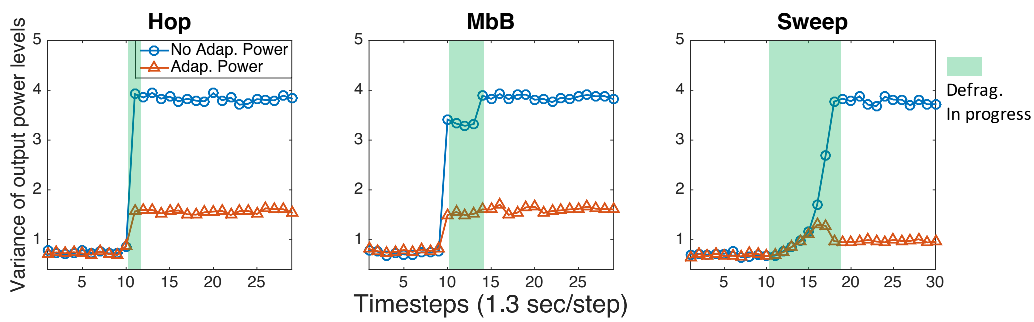

Fig. 2. Differences between power variances during three defragmentation processes when applying adaptive power adjustments in Experiment 1

Experiment 2, shown in Fig. 3, tests a randomly fragmented spectrum scenario with 13 ON channels at Ch. 4, 5, 8, 10, 12, 13, 16, 18, 19, 20, 21, 23, 24, with a super-channel defined at Ch. 18-20. Both Hop and MbB are applied to directly shift the super-channel to Ch. 1-3, and see a 62% improvement when applying adjustments. For Sweep, the first seven channels are individually swept to respective locations in Ch. 1-7. Then the super-channel is swept as a single entity to Ch. 8-10, followed by the last three channels shifting individually to Ch. 11-13. Thus, we achieve a complete defragmentation of the spectrum. Sweep sees an 89% improvement under adaptive power adjustments.

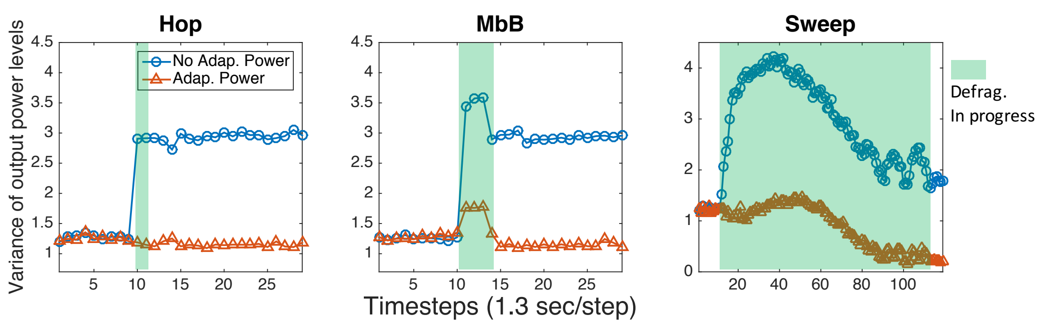

Fig. 3: Differences between power variances during three defragmentation processes when applying adaptive power adjustments in Experiment 2

4. Conclusion

EDFA power excursions present an unsolved challenge for both laboratory and live network operations, and can result in exacerbated post-EDFA power discrepancy during spectral defragmentation processes. We introduce a set of intelligent pre-adjustment strategies enabled by ML that apply to Hop, MbB, and Sweep defragmentation methods. Showcased experimentally, the approach is capable of significant reduction of post-EDFA power discrepancy with adaptive and dynamic adjustments of pre-EDFA sub-channel powers.

5. Acknowledgement

This project is supported by CIAN NSF ERC under grant EEC-0812072 and NSF CNS-1423105. The authors would also like to thank AT&T Labs for their generous donation of the equipment.

6. References

[1] S. Ba et al., "Route Partitioning Scheme for Elastic Optical Networks With Hitless Defragmentation," *JOCN*, **8**(6), p. 356, 2016.
[2] M. Zhang et al., "Spectrum Defragmentation Algorithms for Elastic Optical Networks using Hitless Spectrum Retuning Techniques," OFC, OW3A.4, 2013.
[3] T. Takagi et al., "Disruption minimized spectrum defragmentation in elastic optical path networks that adopt distance adaptive modulation," in ECOC, 2011.
[4] F. Cugini et al., "Push-pull defragmentation without traffic disruption in flexible grid optical networks," JLT, 31(1), p. 125, 2013.
[5] D. C. Kilper et al., "Optical power dynamics in wavelength layer software defined networking," in Advanced Photonics, 2015.
[6] A. S. Ahsan et al., "Excursion-Free Dynamic Wavelength Switching in Excursion-Free Dynamic Wavelength Switching in Amplified Optical Networks," JOCN, 7(9), p. 898, 2015.
[7] Y. Huang et al., "A Machine Learning Approach for Dynamic Optical Channel Add / Drop Strategies that Minimize EDFA Power Excursions," in ECOC, 2016.
[8] P. J. Lin, "Reducing Optical Power Variation in Amplified Optical Network," in ICCT, 2003.