Neural Network Based Wavelength Assignment in Optical Switching

Craig Gutterman\textsuperscript{1}, Weiyang Mo\textsuperscript{2}, Shengxiang Zhu\textsuperscript{2}, Yao Li\textsuperscript{2}, Daniel C. Kilper\textsuperscript{2} and Gil Zussman\textsuperscript{1}

\textsuperscript{1}Electrical Engineering, Columbia University
\textsuperscript{2}College of Optical Sciences, University of Arizona

Aug. 21, 2017
Optical Networks

- Modern networks are **static, ‘set and forget’ approach**
  - Overprovisioned, manual configuration

- Optical devices are capable of switching and dynamic operation
  - Wavelength selective switches, modulators, tunable lasers, and filters
Dynamic Optical Networks

• Currently no real-time capacity management, wavelength reconfiguration and service provisioning in optical networks
  – Provision wavelength
  – Re-route/switch wavelength

• Service providers need to ensure reconfiguration will not affect service level agreement
  – Consider optical power dynamics: WDM channels are highly interactive
    • Disturbances grow in cascade, propagate through network
  – Current provisioning takes minutes to days per wavelength—cautious ‘adiabatic’ tuning to avoid disruptions

• Goal: Given active channel conditions, determine the placement of a new wavelength channel to minimize power excursions under real-time, rapid reconfiguration
Optical Network Components

- **Erbium-Doped Fiber Amplifier (EDFA):** Optical amplifier to boost intensity of optical signal
- **Reconfigurable Optical Add-Drop Multiplexer (ROADM):** Allows for remotely switching traffic at wavelength layer
- **Wavelength Selective Switch (WSS):** Routes signals between optical fibers on a per-wavelength basis
- **Variable Optic Attenuator (VOA):** Reduces power level of optical signal
- **Optical Channel Monitor (OCM):** Used to measure channel power
Power Excursions

- Erbium-Doped Fiber Amplifier (EDFAs)
  - Use Automatic Gain Control (AGC) to maintain constant target mean gain
  - Amplify input channel power to target gain
  - Results in deviations that perturb active channels causing excursions

![Diagram](image)

**Objective:** Minimize $\Delta G$
EDFA Gain Model

\[ P_{\text{out total}} = G_T P_{\text{in total}} \]
\[ P_{\text{out total}} = \sum_{j=1}^{N} P_{\text{in j}} G_j \]
\[ \Delta G = \frac{N G_T + G_T}{N G_T + G_{N+1}} \]

\( \lambda_1 \)
\( \lambda_2 \)

Gain Profile with \( \lambda_1 \)
Gain Profile with \( \lambda_1 \) and \( \lambda_2 \)

- \( G_T \): target mean gain
- \( G_j \): channel gain
- \( P_{\text{in total}} \): total input power
- \( P_{\text{out total}} \): total output power
- \( N \): number of active channels
- \( \Delta G \): power excursion

Target Mean Gain \( G_T \)

Channel Index

Gain (dB)
Previous Work

• Planning phase of dynamic optical network
  – [Angelou et al., 2012], [Doverspike et al., 2012], [Gringeri et al., 2013]
  – Dynamic optical design and architecture
  – No dynamic optical network algorithms for physical layer

• Analytical models
  – [Ishii et al., 2016], [Junio et al., 2012]
  – Non-real time (Junio), poor accuracy (Ishii)

• Machine learning for wavelength assignment
  – [Huang et al., 2017]
  – Only 24 channels, single hop system
  – Kernel Bayesian regression
Our Contributions

• Design an optical testbed
• Collect real world data on optical network performance
• Develop a scalable neural network to predict power excursions at the physical layer
• Evaluate the performance using optical testbed measurements
Optical Testbed and Data Collection
Optical Testbed

- **5-ROADM experiment setup**
  - Nodes separated by 4 standard Single Mode Fiber (SMF) spans
  - Each span has 2 EDFAs
  - 90 channel DWDM source
  - Measurements +/- 0.1 dB accuracy
Optical Testbed – Measurements Collection

2,100 measurement cases
- 3-5 active channels
- 40 new channels added/removed one at a time
- 40 samples/case

84,000 samples
- Training: 1,680 cases (67,200 samples)
- Validation: 210 cases (8,400 samples)
- Test: 210 cases (8,400 samples)
- Data collection: ~4 seconds/sample
2,100 measurement cases

- 3-5 active channels
- 40 new channels added/removed one at a time
- 40 samples/case

84,000 samples

- Training: 1,680 cases (67,200 samples)
- Validation: 210 cases (8,400 samples)
- Test: 210 cases (8,400 samples)
- Data collection: ~4 seconds/sample
Feed forward Neural Network

- Predict the maximum power excursion among all active channels

\[ y^k = \max_{j \in \text{Active Channel}} \Delta G^k_j, \; k \in \text{Available Channel} \]

- Goal: Given initial channel conditions predict which new channel will minimize the effect of power excursion

New Channel \(= k^* = \arg \min_{k \in \text{Available Channels}} \hat{y}^k \)

\[ \bar{x} = [\bar{x}_{\text{Active}}, \bar{x}_{\text{Available}}] \]

\[ = [x_1, \ldots, x_{90}, 0, \ldots, 1, \ldots, 0] \in \{0, 1\}^{180} \]

\[ x_i = \begin{cases} 1, & \text{if } i \in \text{Active Channel}, \; \forall i \in \{1, \ldots, 90\} \\ 0, & \text{Otherwise} \end{cases} \]

\[ x_{i+90} = \begin{cases} 1, & \text{if } i = k^* \\ 0, & \text{Otherwise} \end{cases} \]
Feed forward Neural Network

- Test neural network parameters
  - Activation functions
  - Number of layers
  - Number of neurons per layer

- Optimal Performance
  - 4 Hidden Layers: 180, 120, 30, 15 neurons
  - Tanh activation for hidden layers
  - Linear activation for the output layer

\[
x_i = \begin{cases} 
1, & \text{if } i \in \text{ActiveChannel}, \forall i \in \{1, \cdots, 90\} \\
0, & \text{Otherwise}
\end{cases}
\]

\[
x_{i+90} = \begin{cases} 
1, & \text{if } i = k^* \\
0, & \text{Otherwise}
\end{cases}
\]
Outline

• Design an optical testbed
• Collect real world data on optical network performance
• Develop a scalable neural network to predict power excursions at the physical layer

• Evaluate the performance using optical testbed measurements
  – Prediction
  – Recommendation
  – Classification
Prediction Results

- Training: 67,200 samples
- Test Points: \( n = 8,400 \) samples
- Excursion:

\[
y_l = \max_{j \in \text{Active Channel}} \Delta G_j^k, \quad k \in \text{Available Channel}, \quad \forall l \in \{1, \ldots, n\}
\]

- For Neural Network (NN) comparison Ridge Regression (RR) is trained and parameters tuned on the same data set
Prediction Performance Evaluation

- Training: 67,200 samples
- Test Points: $n = 8,400$ samples
- Excursion:
  \[ y_l = \max_{j \in \text{Active Channel}} \Delta G^k_j, \ k \in \text{Available Channel}, \ \forall l \in \{1, \ldots n\} \]
- Test data Root Mean Square Error (RMSE):
  \[ RMSE_{Test} = \sqrt{\frac{1}{n} \sum_{l=1}^{n} (y_l - \hat{y}_l)^2} \]
- Test data Mean Absolute Error (MAE):
  \[ MAE_{Test} = \frac{1}{n} \sum_{l=1}^{n} |y_l - \hat{y}_l| \]
**ε-Recommendation Accuracy**

- Training: 67,200 samples
- Test Cases: 210
- Test Points: $n = 8,400$ samples

**Recommendation Error**

$$
\text{Recommendation Error} = \epsilon = |y^{\hat{k}}_* - y^{k_*}|
$$

- **Training:** 67,200 samples
- **Test Cases:** 210
- **Test Points:** $n = 8,400$ samples

---

- **Neural Network**
- **Ridge Regression**

---

- **Case Number**
- **ε (dB)**
- **F(ε) (%)**

---

- **Neural Network**
- **Ridge Regression**

---

- **90 Channel DWDM Source**
  - λ = 20
  - λ = 22
  - λ = 67
Classification Accuracy

- Classification: predict if excursion below threshold
  1 : Below Threshold
  0 : Above Threshold
- Set threshold of excursion to 0.5 dB

True Positive Rate (TPR) or Recall = \( \frac{TruePositive}{TruePositive + FalseNegative} \)

False Positive Rate (FPR) = \( \frac{FalsePositive}{FalsePositive + TrueNegative} \)
Classification: Precision at Recall Rate (PSRR)

• Priority for reliable optical system operation:
  1. Minimize false positives: Violates SLA
  2. Minimize false negatives: Wastes open bandwidth

True Positive Rate (TPR) or Recall =
\[
\frac{TruePositive}{TruePositive + FalseNegative}
\]

Precision (PPV) =
\[
\frac{TruePositive}{TruePositive + FalsePositive}
\]
Summary

• Fabricated a 90-channel multi-hop ROADM optical testbed for data collection

• Developed a neural network to predict optical power excursions

• Demonstrated very good performance
  – Prediction: Average prediction error on each channel below 0.1 dB
  – Recommendation: Pick best channel (out of 40), over 70% of the time
  – Classification: Predict maximum excursion to be less than 0.5 dB threshold with a precision of over 99% while obtaining a true positive rate greater than 55%

• Future Work:
  – Machine learning solution for ring and mesh networks
  – Implementation in Optical Software Defined Networks (OSDN)
Thank you!

clg2168@columbia.edu

Craig Gutterman\textsuperscript{1}, Weiyang Mo\textsuperscript{2}, Shengxiang Zhu\textsuperscript{2}, Yao Li\textsuperscript{2}, Daniel C. Kilper\textsuperscript{2} and Gil Zussman\textsuperscript{1}

Neural Network Based Wavelength Assignment in Optical Switching