POLITECNICO
MILANO 1863

Workshop: AI on Optical Networks @BUPT

Emerging Research Directions for Machine Learning in Optical Networks

Speaker: Massimo Tornatore
Associate Professor, Politecnico di Milano, Italy

Beijing University of Posts and Telecommunications
Nov. 2nd, 2018, Beijing, China

What is Machine Learning?

- *"Field of study that gives computers the ability to learn <u>without being explicitly programmed"</u> (A. Samuel, 1959)*
- *<u>"… through data observation</u>"*

- <u>For our purposes</u>: An set of math/statistical tools to make predictions/decisions based on monitored data

  …in the context of optical networks

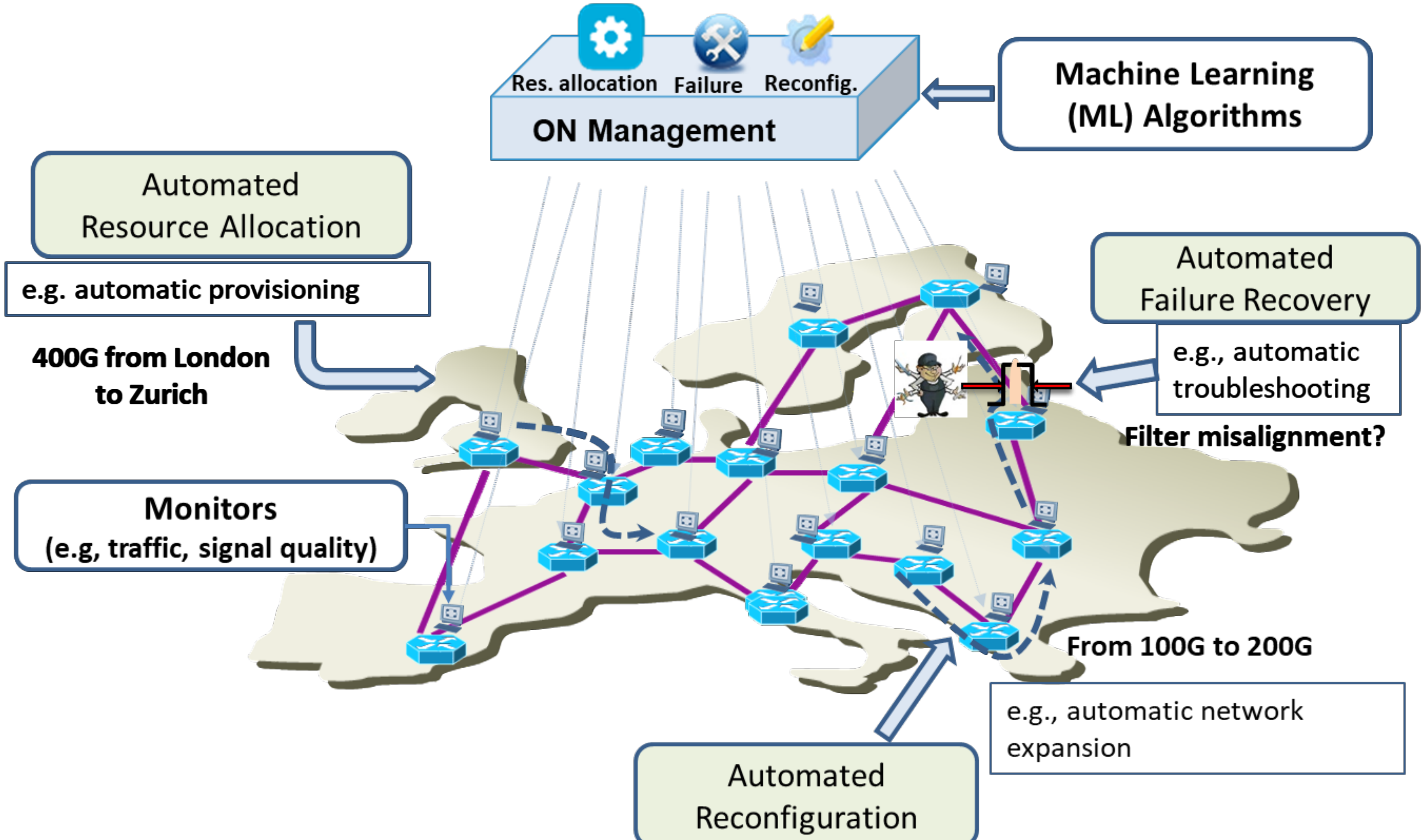
- Confusing overlap with other terms: Artificial Intelligence, Deep Learning, Data Analytics, Data Mining, etc.

Why only now in optical networks?

- Dominating complexity
  - Coherent Trasmission /Elastic Networks
    - Several system parameters: channel bandwidth, modulation formats, coding rates, symbol rates..
- Lack of skilled workforce
  - NTT warning *(OFC 2017)*: aging population, increasing competition for young STEM workforce
- 5G Transport
- New enablers @ *Mngt&Cntr* plane
  - Software Defined Networking
  - Edge computing
  - OPM's (some are for free.. as in coherent receivers..)

Automation of Optical Network Management

- Management is still largely manual/human-based!



Covered topics

- QoT estimation and Routing and Spectrum Assignment
- Soft-Failure Mode Identification
- Quickly, some other applications…

I'll share my experience in developing ML-based solutions in Optical Networks

Motivation
Increasing «degrees of freedom»

- A wider range of **degrees of freedom** (parameters) is available to system engineers:
  - path
  - spectrum
  - modulation format
  - baud rate
  - FEC coding
  - single/multicarrier transmission
  - nonlinearity mitigation solution
  - adaptive channel spacing
  - …
- Combinations of these lighpath parameters grow dramatically
- Possibly, for all of these combinations, we shall calculate a QoT

Existing (pre-deployment) estimation techniques for lightpath QoT

- **“Exact” analytical models** estimating physical layer impairments (e.g., split-step Fourier method…)
  - 🙂 Accurate results
  - ☹ Heavy computational requirements
  - ☹ Not scalable to large networks and real time estimations

- **Marginated formulas** (Power Budget, Gaussian model…)
  - 🙂 Faster and more scalable
  - ☹ Inaccurate, high margination, underutilization of network resources (up to extra 2 dB for design margins [1])

[1] Y. Pointurier, "Design of low-margin optical networks," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 9, no. 1, pp. A9-A17, Jan. 2017. doi: 10.1364/JOCN.9.0000A9

Machine Learning as an alternative approach?

- Machine Learning exploits knowledge extracted from field data…
  - QoT of already established lightpaths, e.g. using monitors at the receiver
- …. to predict the QoT of unestablished lightpaths

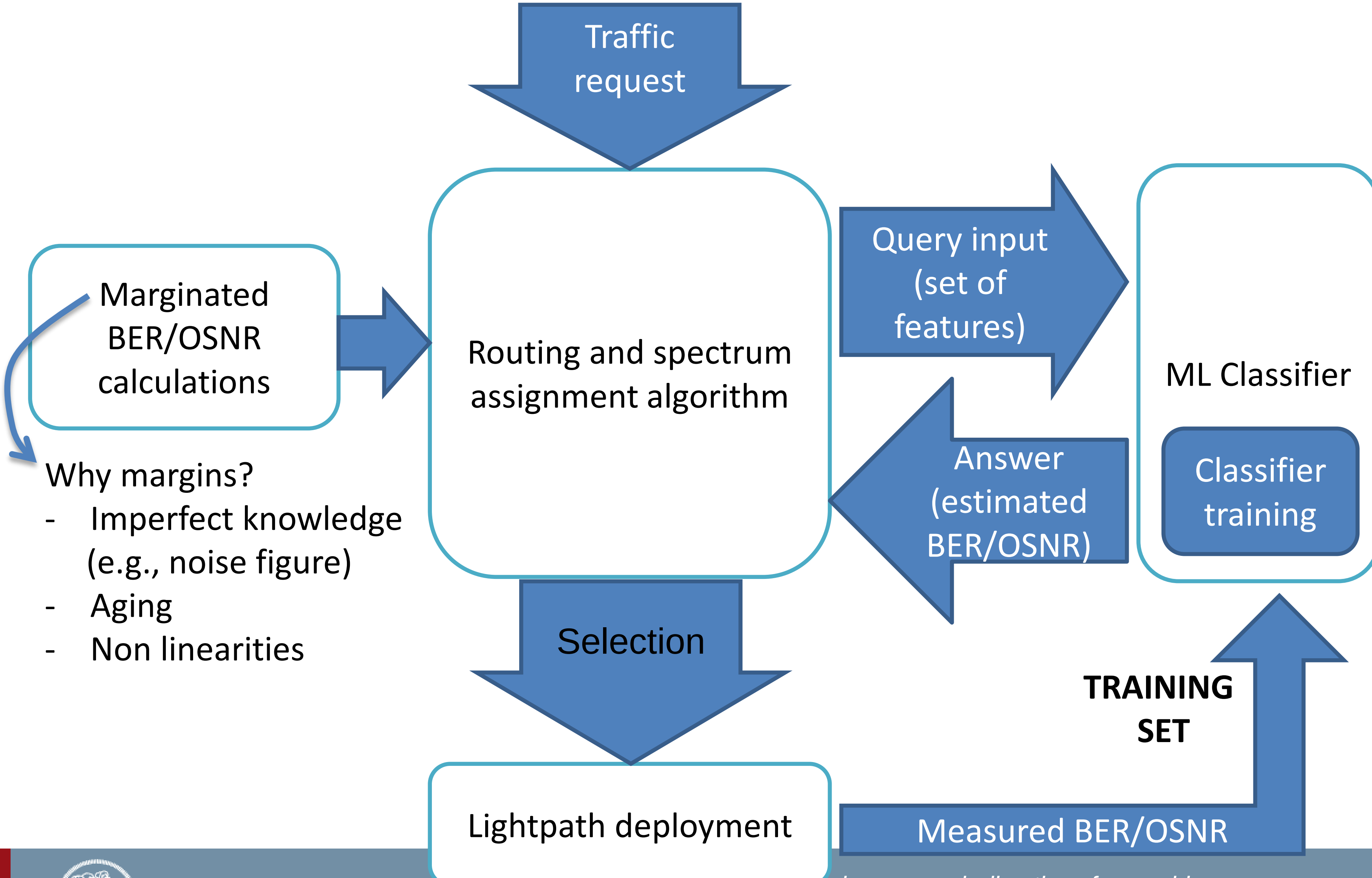
  - No need for complex analytical models
  - Fast and scalable
  - Requires training phase with historical data

    How long must the training phase be?

    How accurate will the estimation be?

    Objectives of our numerical analysis….

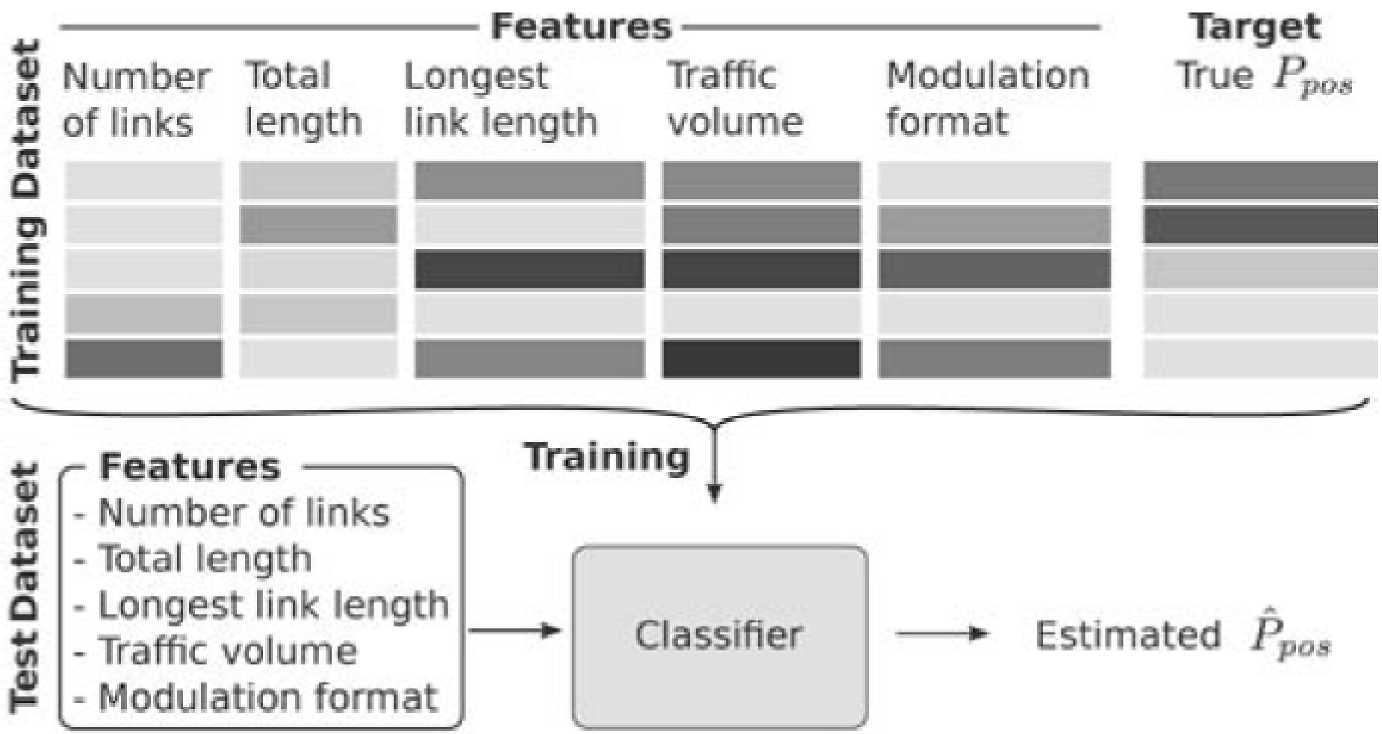
RSA interplays with QoT estimation



How does it work?

A possible implementation of ML-based QoT estimation

- Input: set of lightpath features
- Output: probabilty that BER ≤ T*



(Case of **local knowledge**, but we can add more features for **network knowledge**)

C. Rottondi, L. Barletta, A. Giusti and M. Tornatore, *A Machine Learning Method for Quality of Transmission Estimation of Unestablished Lightpaths*, JOCN2018

How our proposed ML classifier works
Case 2

- To the previous 6 feature we add, for the «most interfering left and right neighbors»:
  - guardband
  - traffic volume
  - modulation format

(Case of **complete** knowledge)

- Note: these additional six features are chosen with the intent to capture cross-channel nonlinear effects

Which Machine Learning Technique?

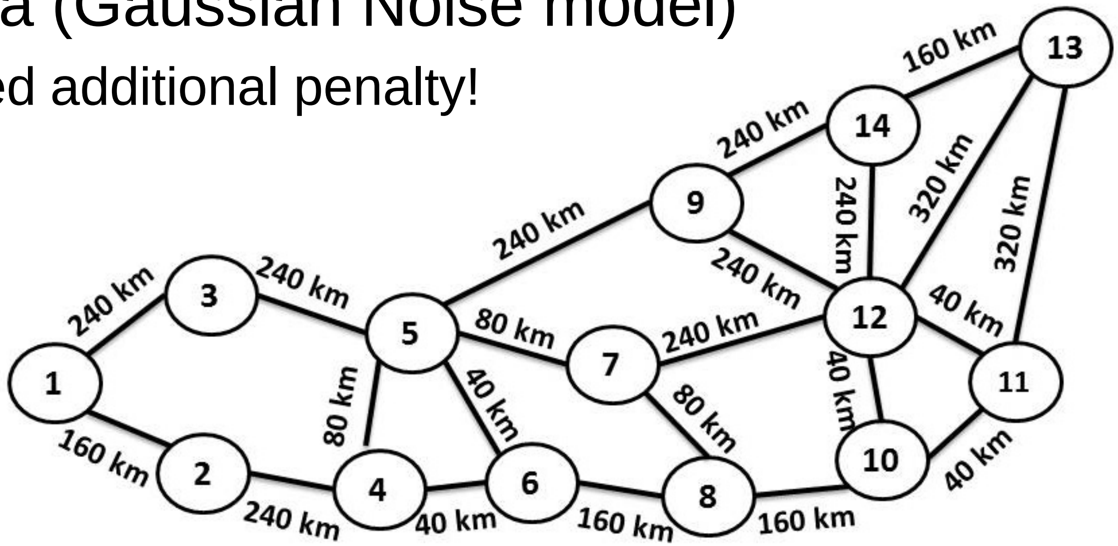
- We use a Random Forest (RF) classifier with 25 estimators
- To take this choice, we have compared:
  - 5 RF classifiers
  - 3 k-Nearest-Neighbor classifiers

Algorithm	Training time (s)	Test time (s)	AUC	Accuracy
Dummy classifier	0.048979	3.83 e-07	0.501	0.539
1 Nearest Neighbor	1.183121	4.83 e-05	0.959	0.957
5 Nearest Neighbor	1.085116	5.05 e-05	0.991	0.965
25 Nearest Neighbor	1.211694	6.91 e-05	0.996	0.965
Random Forest 1 tree	0.076944	3.96 e-07	0.991	0.965
Random Forest 5 trees	0.180835	6.24 e-07	0.995	0.970
Random Forest 25 trees	0.721042	1.56 e-06	0.996	0.968
Random Forest 100 trees	2.830545	5.32 e-06	0.996	0.966
Random Forest 500 trees	14.052182	2.63 e-05	0.996	0.966

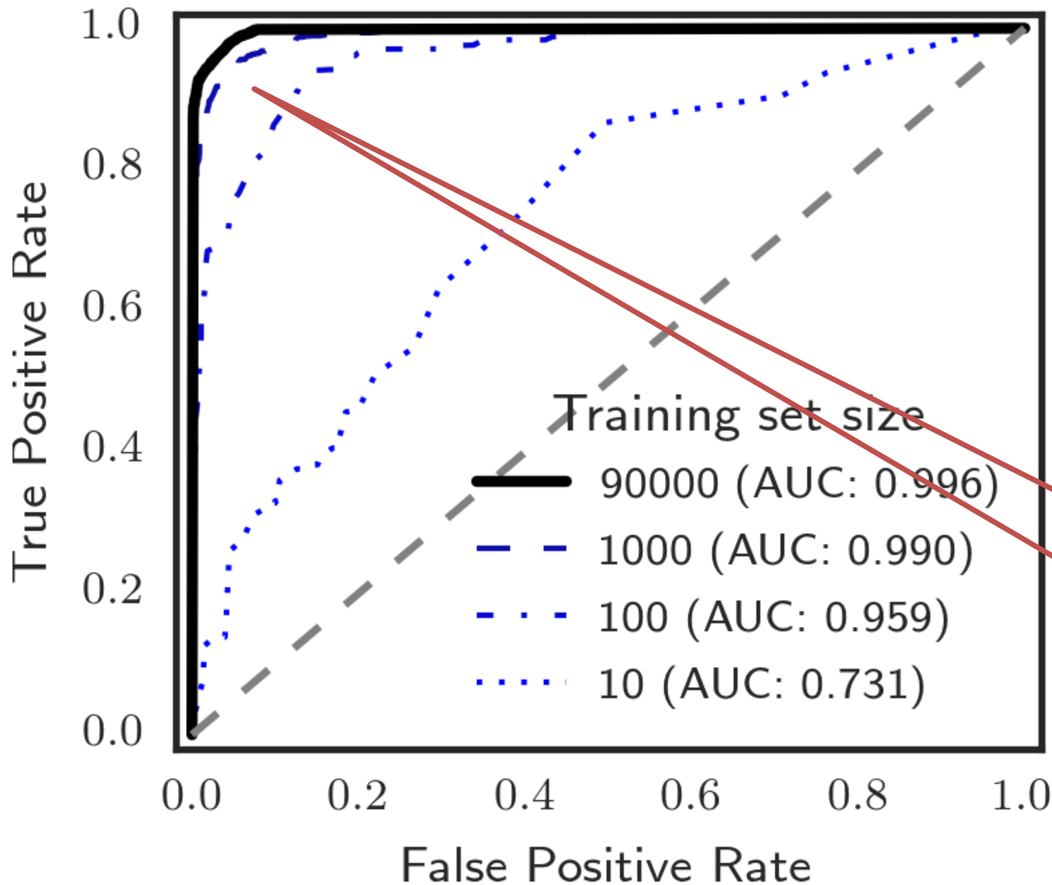
- RF with 25 estimators provided the best trade-off between performance and computational time

Training and Testing Scenario

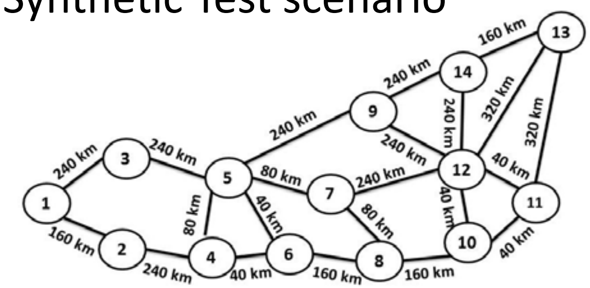
- Japanese optical network
- Flexgrid @ 12.5 GHz slices
- Transceivers @ 28 GBaud with adaptive modulation formats
  - DP-BPSK, -QPSK, -8-QAM, -16-QAM, -32-QAM, -64-QAM
- Traffic requests: [50;1000] Gbps
- Synthetic training data (Gaussian Noise model)
  - With expneg distributed additional penalty!



How big shall training dataset be?



- Synthetic Test scenario



- Accuracy: Area under the ROC curve (AUC)

Take-Away 1: Training phase has a reasonable duration

C. Rottondi, L. Barletta, A. Giusti and M. Tornatore, *A Machine Learning Method for Quality of Transmission Estimation of Unestablished Lightpaths*, JOCN2018

How to build the training dataset?

- Use historical data
  - ☹ We will never observe samples of with too high BER!!
- Use random probes:
  - ☹ Very costly (high spectrum occupation)
- Use selective probes:
  - ☺ Lower spectrum occupation, good accuracy

TABLE V: AUC comparison of probing approaches

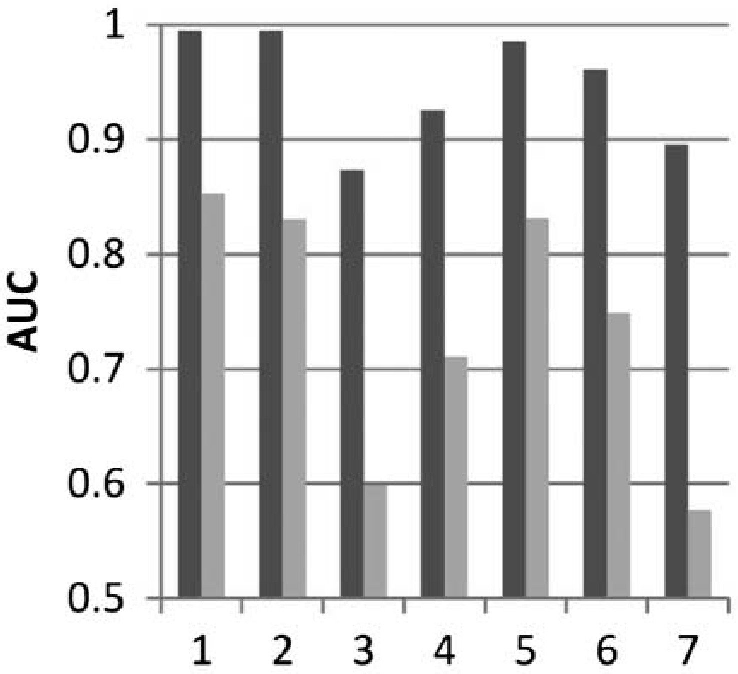
Training set	AUC (full testing dataset)
C (historical)	0.77
C (selective, 5% probes)	0.85
C (selective, 10% probes)	0.87
C (selective, 25% probes)	0.89
C (selective, 50% probes)	0.89
A (random)	0.98

Analysis of feature relevance

- Removing irrelevant «ML-input features» makes the system less costly and less complex to manage

TABLE IV: The considered feature subsets

	S1	S2	S3	S4	S5	S6	S7
number of links	✓	✓	✓	✓			
lighpath length	✓	✓	✓	✓	✓	✓	
length of longest link	✓	✓	✓	✓			
traffic volume	✓	✓	✓		✓		✓
modulation format	✓	✓		✓	✓	✓	✓
guardband, modulation format and traffic volume of nearest left and right neighbor	✓						

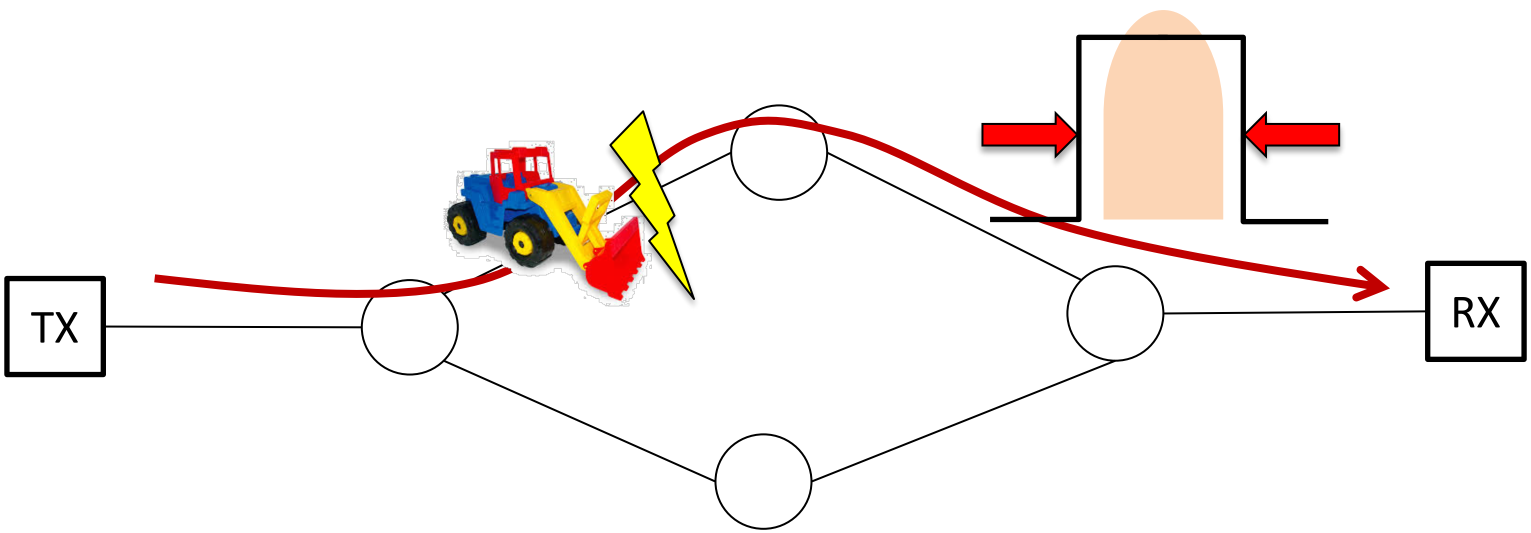


Covered topics

- QoT estimation and Routing and Spectrum Assignment
- Soft-Failure Mode Identification
- Quickly, some other applications…

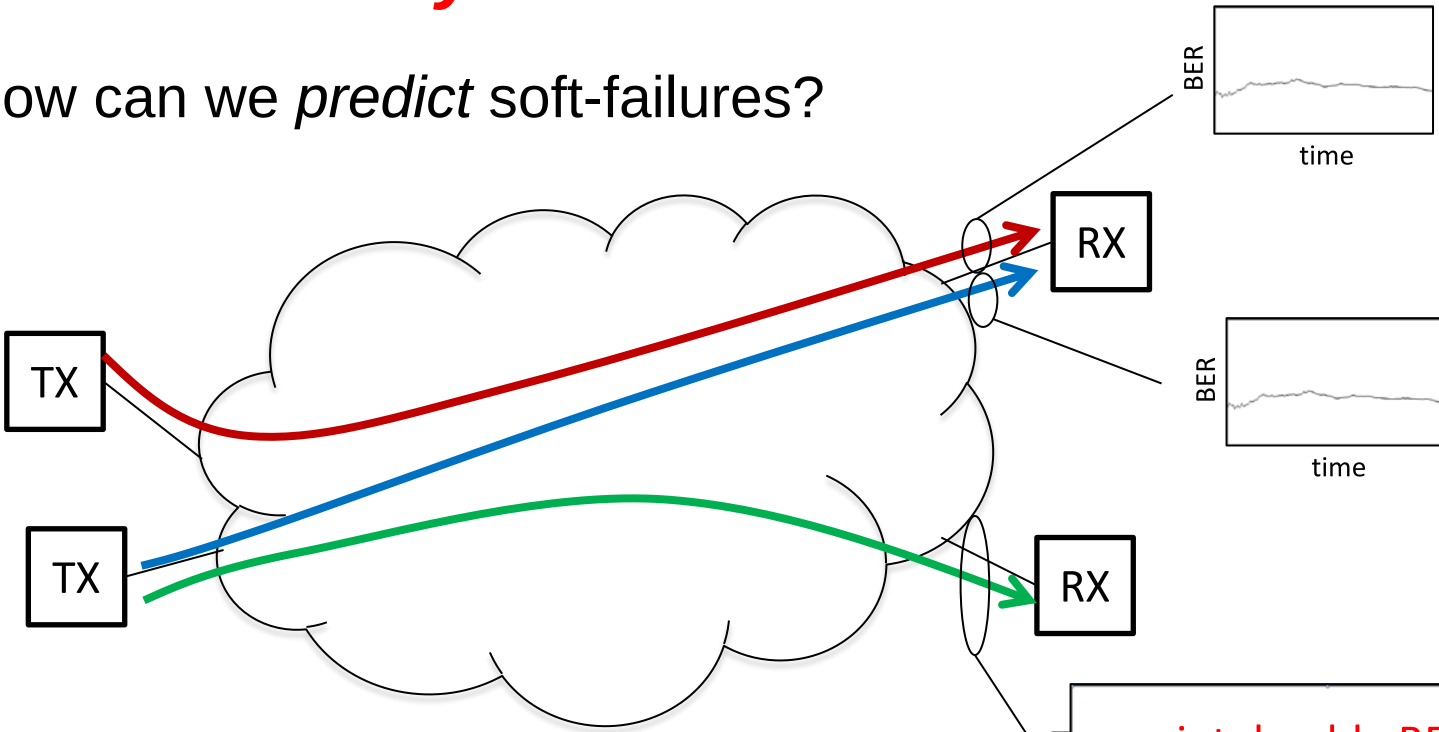
Two main failure types in optical networks

- Hard-failures
  - Sudden events, e.g., fiber cuts, power outages, etc.
  - Unpredictable, require «protection» (*reactive procedures*)
- Soft-failures:
  - Gradual transmission degradation due to equipment malfunctioning, filter shrinking/misalignment…
  - Trigger early network reconfiguration (*proactive* procedures)

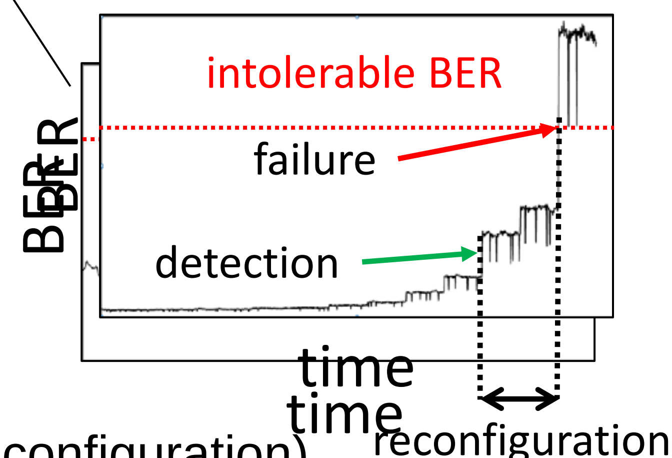


Soft-failure *early detection*

- How can we *predict* soft-failures?



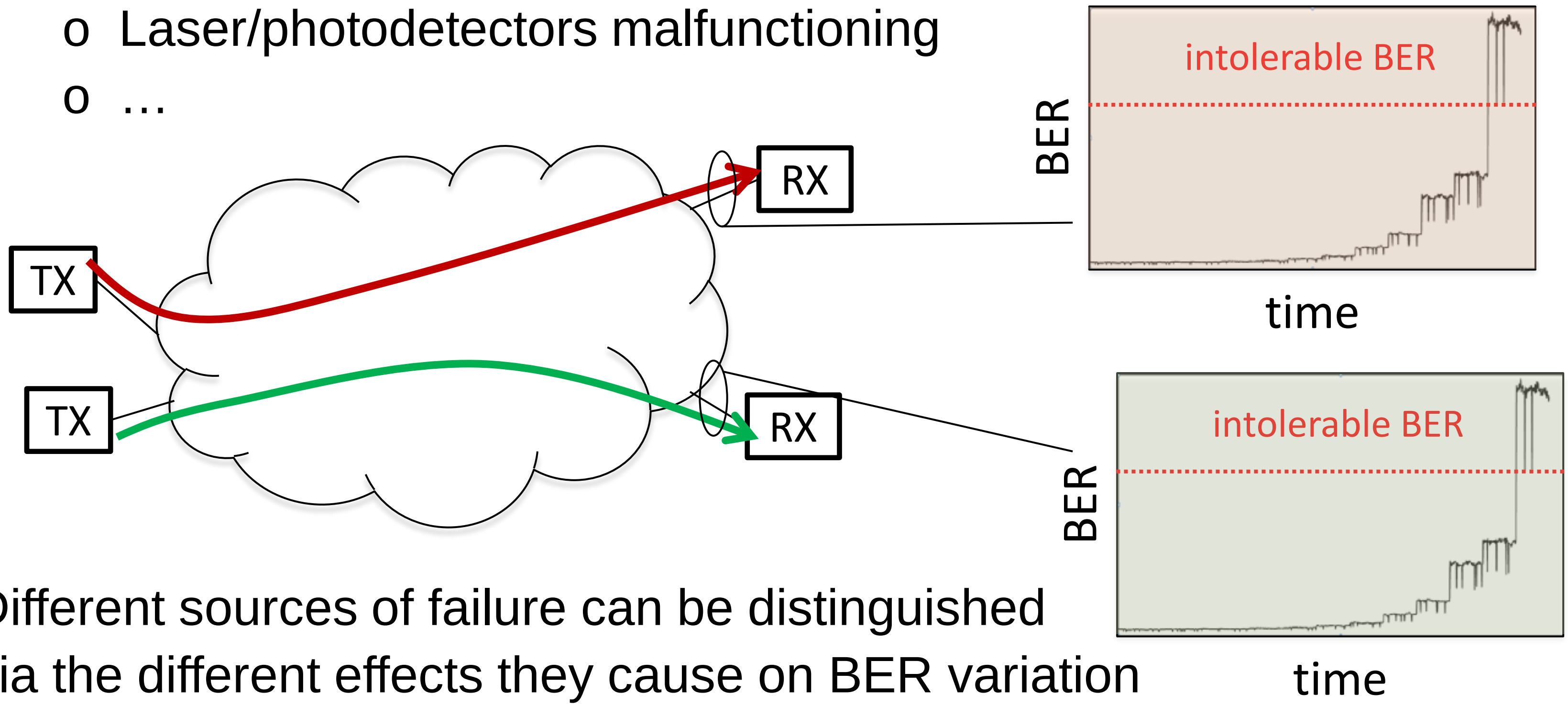
Perform continuous monitoring of
Bit Error Rate (BER) at the receiver…
… until some "anomalies" are detected



Early-detection helps **preventing** service disruption (e.g., through proactive network reconfiguration)

Soft-failure *mode identification*

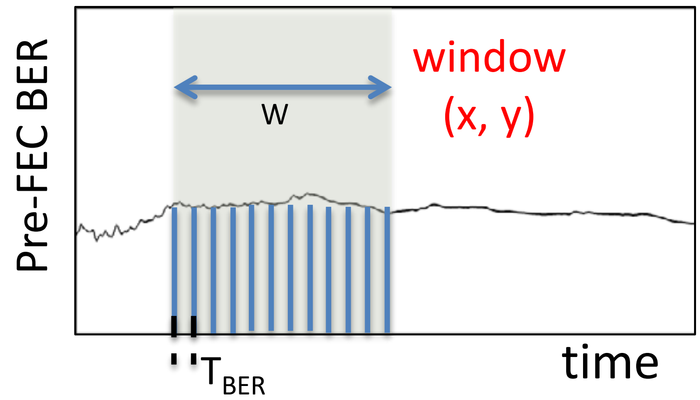
- How can we identify the *mode* of the failure?
  - Failures can be caused by different sources
    - Filters shrinking/misalignment
    - Excessive attenuation (e.g., due to amplifier malfunctioning)
    - Laser/photodetectors malfunctioning
    - …



Different sources of failure can be distinguished via the different effects they cause on BER variation (i.e., via different BER “features”)

2nd Phase of our study

Deciding ML algorithm, Training &Validation



1. Data Retrieval

3 decisions

Validation (optimization of hyperparameters)

BER window → Features → ML algorithm

BER window

Select:
- BER sampling time (T_{BER})
- window size (duration of observation)

Features

- BER statistics:
  - mean
  - min/max
  - standard dev.
- Window spectral components
- Feature Scaling

ML algorithm

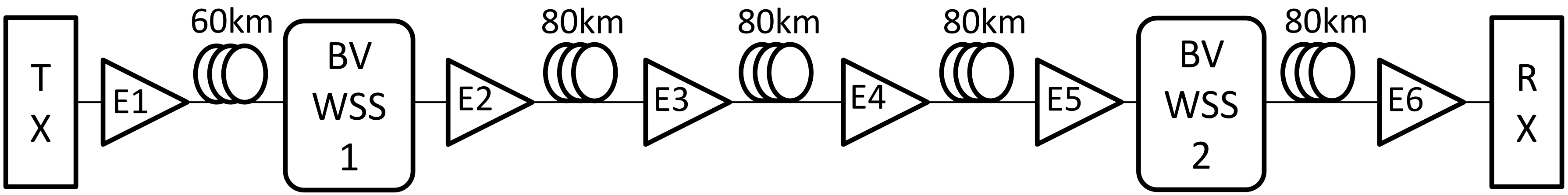
Fault detection:
- Binary SVM
- Random Forest
- Multiclass SVM
- Neural Network

Fault identification:
- Neural Network

3. Prediction and Evaluation

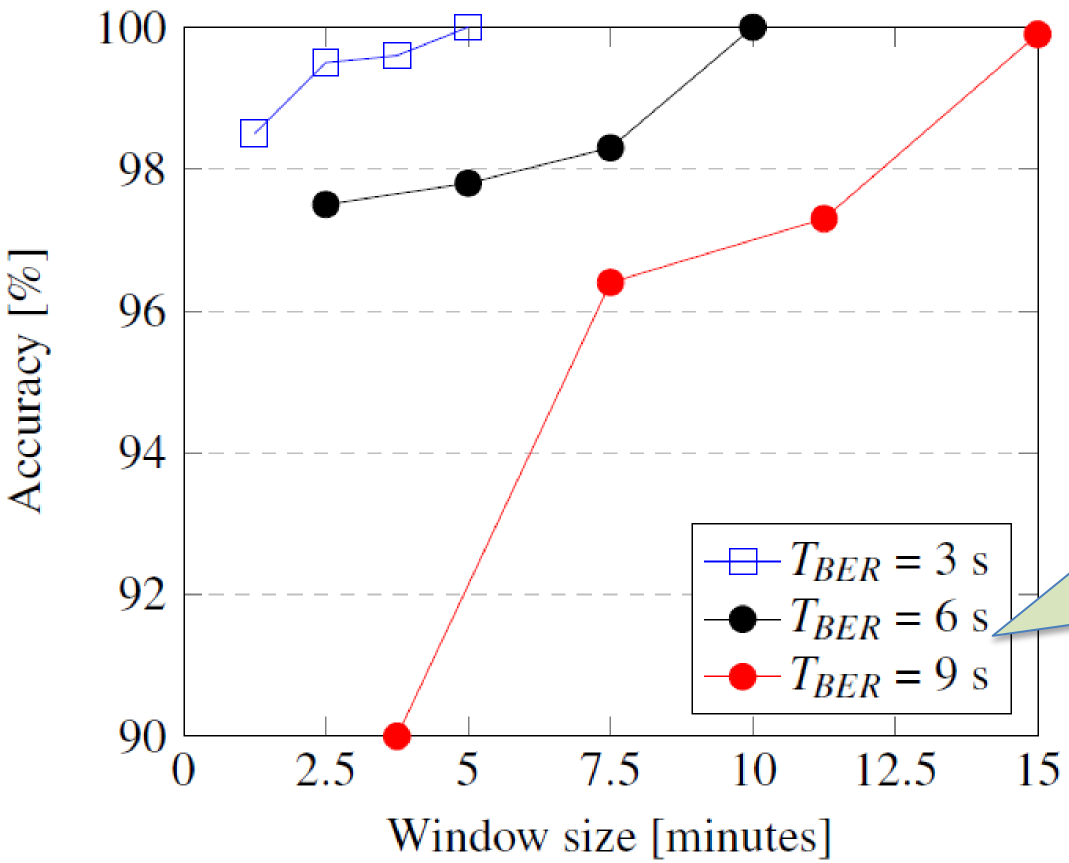
Testbed setup

- Testbed for real BER traces
  - Ericsson 380 km transmission system
    - 24 hours BER monitoring
    - 3 seconds sampling interval
  - PM-QPSK modulation @ 100Gb/s
  - 6 Erbium Doped Fiber Amplifiers (EDFA) followed by Variable Optical Attenuators (VOAs)
  - Bandwidth-Variable Wavelength Selective Switch (BV-WSS) is used to emulate 2 types of BER degradation:
    - **Filter misalignment**
    - Additional attenuation in intermediate span (e.g., due to **EDFA gain-reduction**)



Numerical results: *Identification*
Accuracy vs window features

- Neural Network



Take-away 3: To perform failure-cause identification, fine sampling period of BER is needed

Benefits for operators

- Reduced Time To Repair (TTR)
  - Almost instantaneous troubleshooting
  - TTR from hours/days to minutes/hours?
- Reduced Service Downtime
  - Early detection eliminates a class of failure

- First demonstrations

Vela et al., “BER degradation Detection and Failure Identification in Elastic Optical Networks”, in Journal of Lightwave Technology, vol. 35, no. 21, pp. 4595-4604, Nov. 2017

S. Shahkarami, F. Musumeci, F. Cugini, M. Tornatore, “Machine-Learning-Based Soft-Failure Detection and Identification in Optical Networks,“ in Proceedings, OFC 2018, San Diego (CA), Usa, Mar. 11-15, 2018

Many open questions/challenges!

- **[QoT]** Optical network is a living network
  - Continuos training.. How?
- **[QoT]** How to build the right training set?
  - Rare occurences of false positives -> Low accuracy…
  - Selective probes?
- **[Failure]** What if completely new/unclassified failure arise?
  - «Novelty detection» ?

Overview of other applications

- Physical layer
  1. Optical amplifier control
  2. Modulation format recognition
  3. Nonlinearities mitigation

- Network layer
  1. Traffic prediction and virtual topology design
  2. Flow classification

Classification taken from: F. Musumeci et al., "A Survey on Application of Machine Learning Techniques in Optical Networks", Accepted to IEEE Communication Surveys and Tutorial, available online (Arxiv)

Thanks for your attention!

massimo.tornatore@polimi.it

Metro Haul
H2020

European Commission
Horizon 2020
European Union funding
for Research & Innovation