Keynote Talk
@NIST
Denver, August 2, 2019

POLITECNICO DI MILANO

Machine Learning Applications in Optical Transport Networks

Massimo Tornatore

Politecnico di Milano, Italy & University of California, Davis

@MaxTornatore

Covered topics

- The presentation is organized into two main parts
- Part 1: overview on Machine Learning
  - Basic concepts (supervised/unsupervised learning, neural networks, etc.)
  - Some algorithms
    - Linear regression
    - Neural Networks

Note: The objective is to show **how we applied** ML to **our** research problems

- Part 2: applications of ML to optical-network problems
  - Part 2a): QoT estimation and RSA
  - Part 2b): Failure management
  - Part 2c): Other application at physical and network layer
    - Traffic prediction, virtual topology design,…

What is Machine Learning?

- *“Field of study that gives computers the ability to learn without being explicitly programmed” (A. Samuel, 1959)*
- *“Teaching a computer to automatically learn concepts through data observation”*
- *…*

- For our purposes: An math/statistical instrument to make decisions by inferring statistical properties of monitored data

  …in the context of optical networks

- Sometimes confused with other terms: AI, 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..

- New enablers @ *Mngt&Cntr* plane
  - Software Defined Networking
  - Edge computing
  - OPM's (some of them are for free.. as in coherent receivers..)

- Lack of skilled workforce
  - NTT warning *(OFC 2017)*: aging population, increasing competition for young STEM workforce

Main categories of ML algorithms (1)

- Supervised-learning algorithms
  - We are given "labeled" data (i.e., "ground truth")
  - Main objective: given a set of "historical" input(s) predict an output
    - Regression: output value is continuous
    - Classification: output value is discrete or "categorical"

- An example: Traffic forecasts
  - Given traffic during last week/month/year
    - Predict traffic for the next period (regression)
    - Predict if available resources will be sufficient (classification)

- Other examples
  - Speech/image recognition
  - Spam classifier
  - House prices prediction/estimation

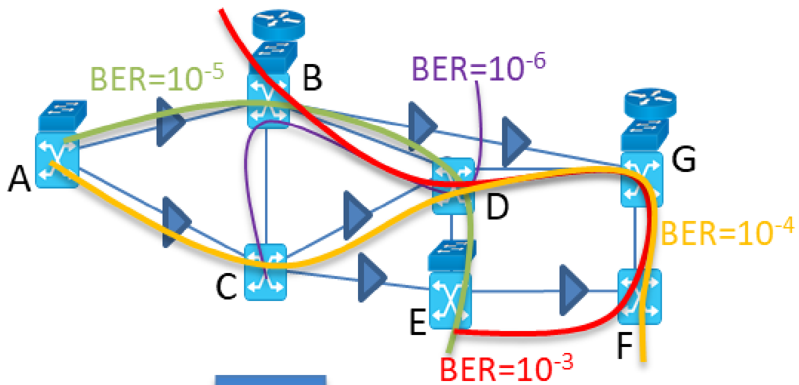
Supervised learning: an «optical» example

Training Phase

λ=1550, path= nodes A-C-D-G-F,
Mod = QPSK, ➔ BER=10^{-5}

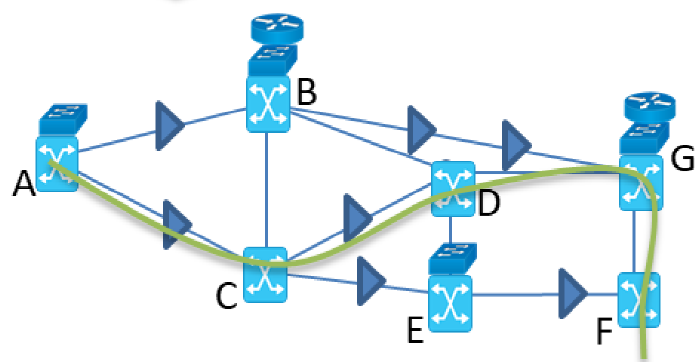
λ=1553, path= nodes B-G-D-F-E,
Mod = QPSK, ➔ BER=10^{-2}

...



Active/Test Phase

Create path: λ= 1553, nodes A-C-D-G-F,
Mod QPSK ➔ BER=?

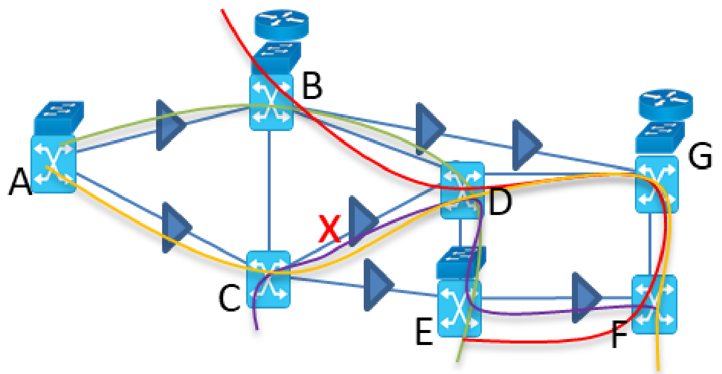


Courtesy of Marco Ruffini and Irene Macaluso

Main categories of ML algorithms (2)

- Unsupervised-learning algorithms
  - Available data is not “labeled”
  - Main objective: derive structures (patterns) from available data
    - Clustering finding “groups” of similar data
    - Anomaly detection
- An example: cell-traffic classification
  - Given traffic traces
  - understand if some cells provide similar patterns
    - Residential, business, close to theatre, cinema, stadium…
    - This information can be used to make network resources planning
- Other example
  - Group people according to their interests to improve advertisement

Unsupervised learning: an optical example



Data:

λ=1550, path= nodes A-B-D-E, Mod = QPSK, BER=10^{-6}

λ=1545, path= nodes B-D-G-F-E, Mod = 16-QAM, BER=10^{-7}

λ=1553, path= nodes A-C-D-G-F, Mod = BPSK, BER=10^{-2}

λ=1544, path= nodes C-D-E-F, Mod = DPQPSK, BER=10^{-2}

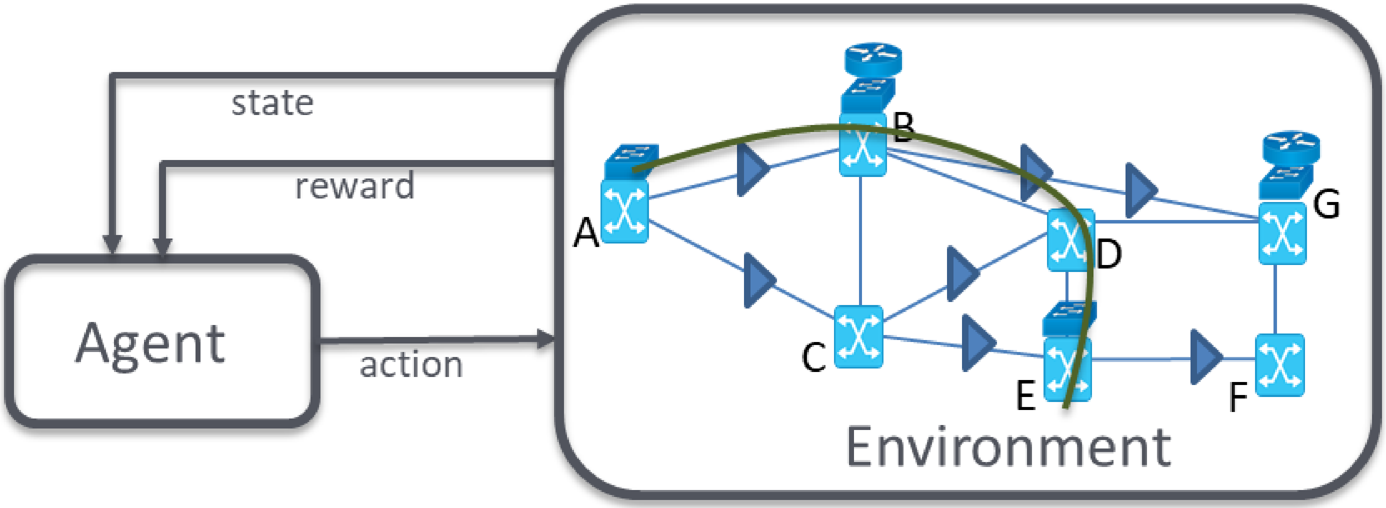
Anomaly detection for link C-D

Courtesy of Marco Ruffini and Irene Macaluso

Main categories of ML algorithms
Cont'd

- Semi-Supervised learning
  - Hybrid of previous two categories
  - Main objective: most of the training samples are unlabeled, only few are labeled
    - Common when labeled data are scarce or expensive
  - Self-training: start with labeled data, then label unlabeled data based on first phase
- Reinforcement learning
  - Available data is not "labeled"
  - Main objective: learn a policy, i.e., a mapping between in inputs/states and actions. Behavior is refined through rewards
  - Methodologically similar to «optimal control theory» or «dynamic programming»
  - Q-learning

Reinforcement learning: an example

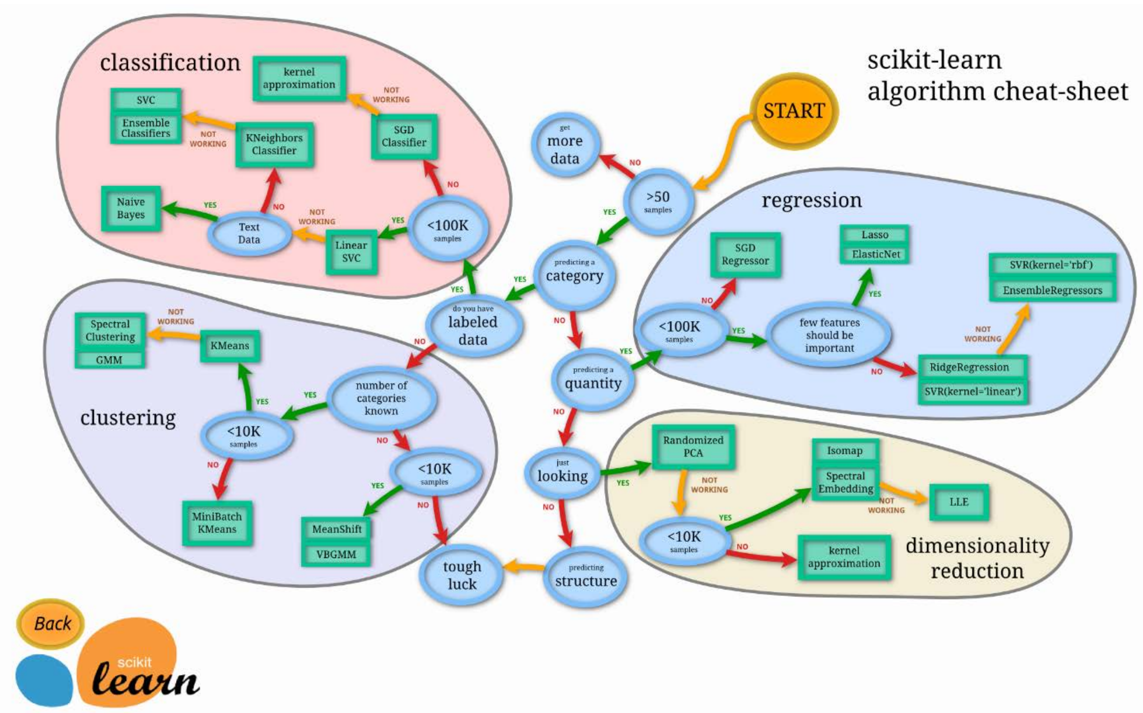


Initial state	Action	State	Reward
λ=1550nm, nodes A-B-D-E, Mod QPSK, BER=10^{-3}	No Change	BER= 10^{-3}	0
λ=1550nm, nodes A-B-D-E, Mod QPSK, BER=10^{-3}	Change: output power channel +5 dBm	BER= 10^{-2}	-1
λ=1550nm, nodes A-B-D-E, Mod QPSK, BER=10^{-3}	Change: Mod BPSK	BER= 10^{-4}	+1

Courtesy of Marco Ruffini and Irene Macaluso

Some algorithms

- Supervised
  - Parametric
    - Linear and logistic regression
    - Neural Networks
    - ..
  - Non parametric
    - K-nearest neighbor
    - Random Forest
    - ...
- Unsupervised
  - Clustering
    - K-means
    - Gaussian Mixture Models
    - …



These slides are not NIST's slides. NIST is not responsible for the content of these slides.

Basic intuition behind neural networks

- If we know the basic charateristics of relation between in input and outputs, math gives us lot of tools:
  - Regression
    - Linear, quadratic, logistic, multivariate, polynomial..

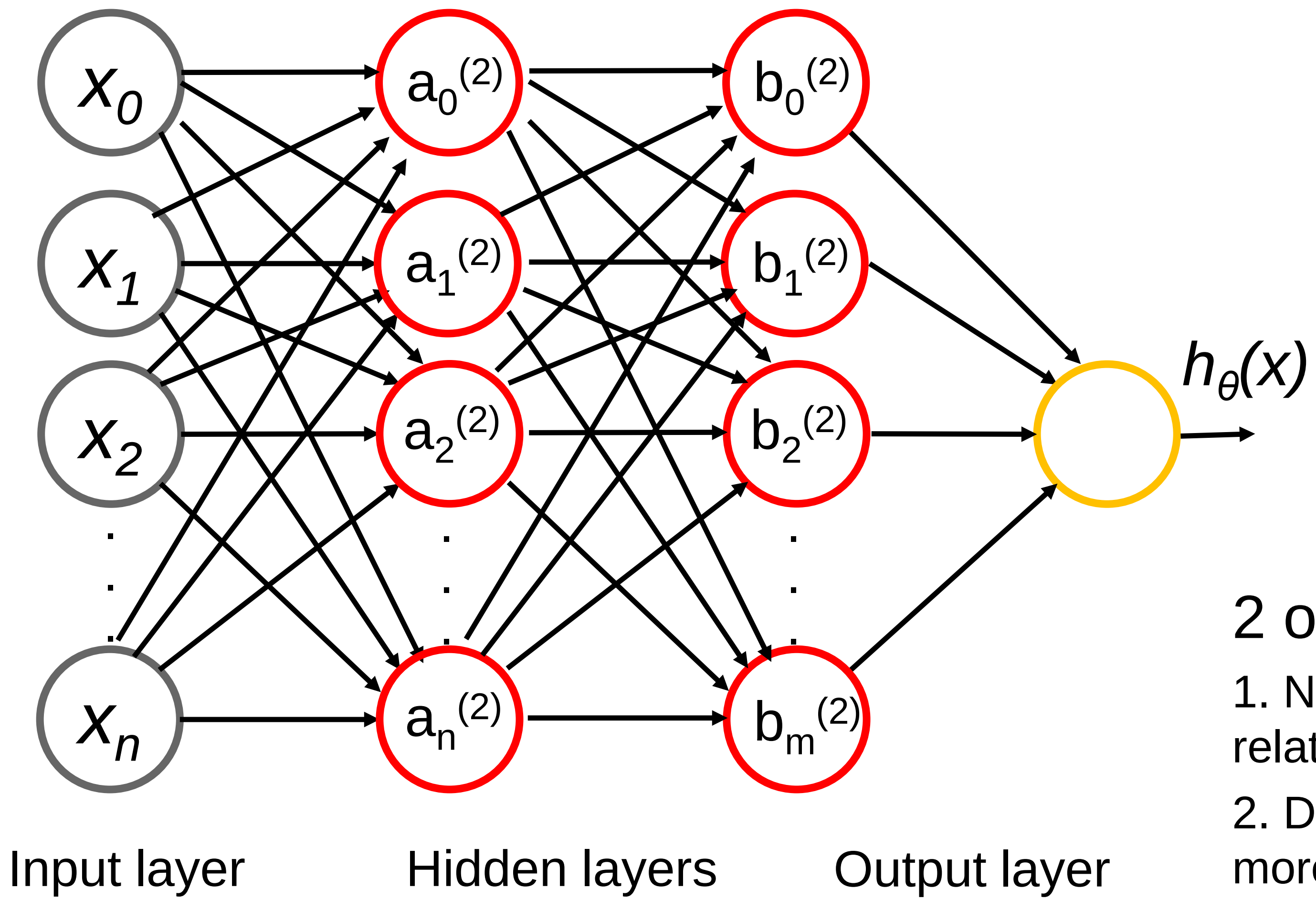
What if the relation is completely unknown?

What if I cannot make any assumption regarding input-output relation?

Neural networks!

Neural Networks (NN) representation

A "collection" of interacting neurons



2 observations:

1. NN can capture any relation between x and y
2. Deep Learning: the more layers, the less decisions shall be taken by a programmer (less «feature engineering)

Outline: Part 2 Applications

1. ML for QoT Estimation for Unestablished Lighpaths
   - C. Rottondi, L. Barletta, A. Giusti and M. Tornatore, *A Machine Learning Method for QoT Estimation of Unestablished Lightpaths*, in IEEE/OSA Journal of Optical Comm.& Netw. Vol. 10, No. 2, Feb. 2018

2. ML for Failure Management
   - Francesco Musumeci ,et al., “A Tutorial on Machine Learning for Failure Management in Optical Networks”, in IEEE/OSA Journal of Lightwave Technology, available online

3. An overview of other applications at network layer
   - F. Musumeci et al., “A Survey on Application of Machine Learning Techniques in Optical Networks”, Submitted to IEEE communication surveys and tutorials, 2019
   - Javier Mata, et a., Artificial intelligence (AI) methods in optical networks: A comprehensive survey, Optical Switching and Networking, Volume 28, 2018, pp. 43-57

Outline: Part 2 Applications

1. ML for QoT Estimation for Unestablished Lighpaths

- C. Rottondi, L. Barletta, A. Giusti and M. Tornatore, *A Machine Learning Method for QoT Estimation of Unestablished Lightpaths*, in IEEE/OSA Journal of Optical Comm.& Netw. Vol. 10, No. 2, Feb. 2018

2. ML for Failure Management

- Francesco Musumeci ,et al., “A Tutorial on Machine Learning for Failure Management in Optical Networks”, in IEEE/OSA Journal of Lightwave Technology, available online

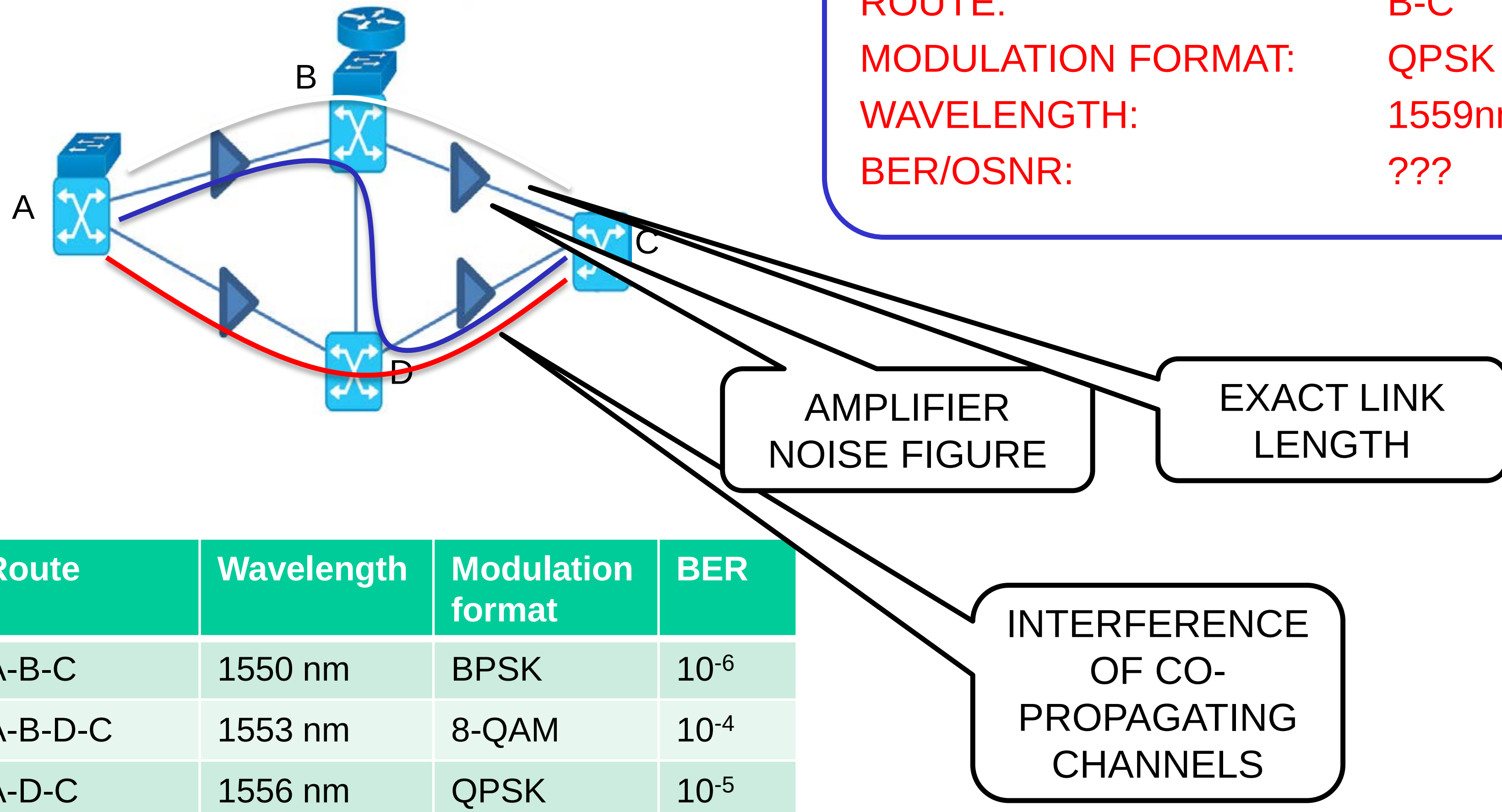
3. An overview of other applications at network layer

- F. Musumeci et al., “A Survey on Application of Machine Learning Techniques in Optical Networks”, Submitted to IEEE communication surveys and tutorials, 2019
- Javier Mata, et a., Artificial intelligence (AI) methods in optical networks: A comprehensive survey, Optical Switching and Networking, Volume 28, 2018, pp. 43-57

Why QoT estimation?

NEW TRAFFIC REQUEST:

ROUTE:	B-C
MODULATION FORMAT:	QPSK
WAVELENGTH:	1559nm
BER/OSNR:	???



Route	Wavelength	Modulation format	BER
A-B-C	1550 nm	BPSK	10^{-6}
A-B-D-C	1553 nm	8-QAM	10^{-4}
A-D-C	1556 nm	QPSK	10^{-5}

How (pre-deployment) QoT estimation is done

- **"Exact" analytical models** (e.g., split-step Fourier method)
  - 🙂 Accurate results
  - ☹️ Heavy computational requirements → not scalable / not real time

- **Margined formulas** (e.g., AWGN model…)
  - 🙂 Faster and more scalable

$$\frac{1}{\mathrm{OSNR_{tot}}} = \sum_{k=1}^{N_{\mathrm{span}}} \frac{1}{\mathrm{OSNR}^{(k)}_{\mathrm{ASE,Rx}}} + \sum_{k=1}^{N_{\mathrm{span}}} \frac{1}{\mathrm{OSNR}^{(k)}_{\mathrm{NL}}} \qquad \mathrm{OSNR}^{(k)}_{\mathrm{ASE,Rx}} = \frac{P^{(k)}_{\mathrm{Tx}}}{h\nu B_n G^{(k)} F^{(k)}}$$

  - ☹️ Analitically accurate, but suffers from inaccurate parameter knowledge.
  - ☹️ 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* methods have been proposed to
  - estimate QoT of unestablished lightpaths
  - using field data, e.g., monitored BER/OSNR at the receiver

☺ 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….

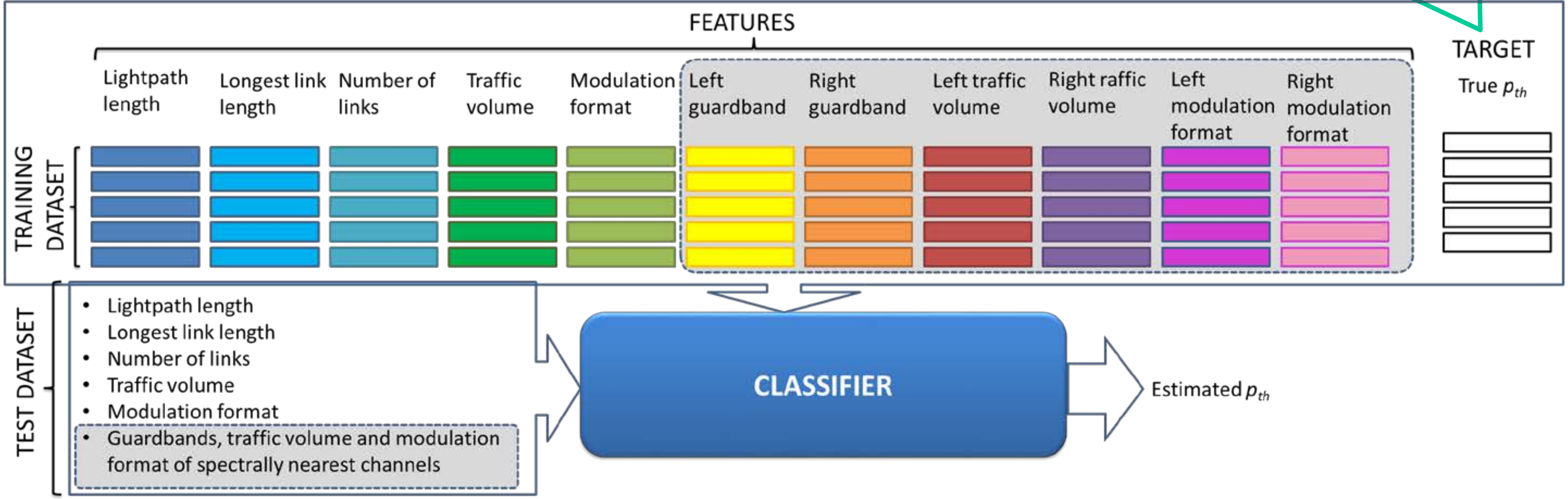
C. Rottondi, L. Barletta, A. Giusti, M. Tornatore "Machine-learning method for quality of transmission prediction of unestablished lightpaths," IEEE/OSA J. of Optical Comm. and Netw., vol. 10, no. 2, pp. A286–A297, Feb 2018.

Proposed ML classifier for QoT estimation

Input: *Lightpath features*

Output: *Prob{BER ≤ T*}*

The classifier is trained on a set of *L* experiments to generate ***ground truth***



These slides are not NIST's slides. NIST is not responsible for the content of these slides.

Which Machine Learning Algorithm?

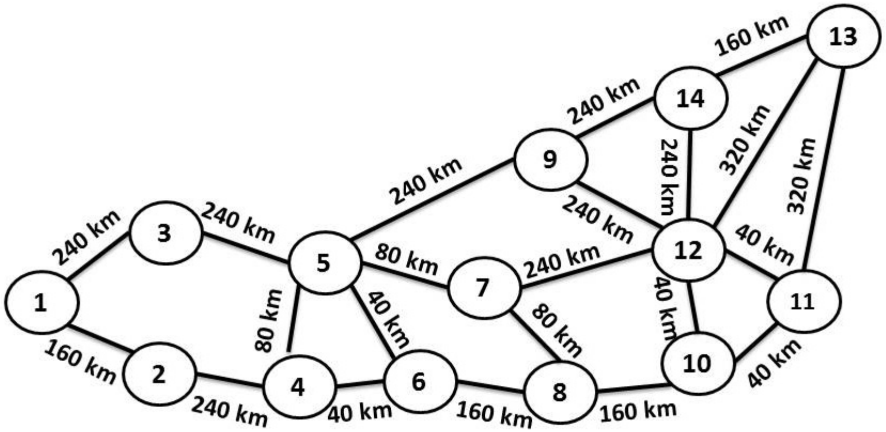
- We used a Random Forest (RF) classifier with 25 estimators
- To take this choice, we compared 5 RFs and 3 kNN classifiers and picked best "accuracy/complexity" tradeoff

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

- But knowledge is rapidly evolving!
  - Neural Networks… SVMs… (parametric approaches)
  - Gaussian processes (return confidence of classification!)

Training and Testing Scenario

- Some results for a Japanese optical network

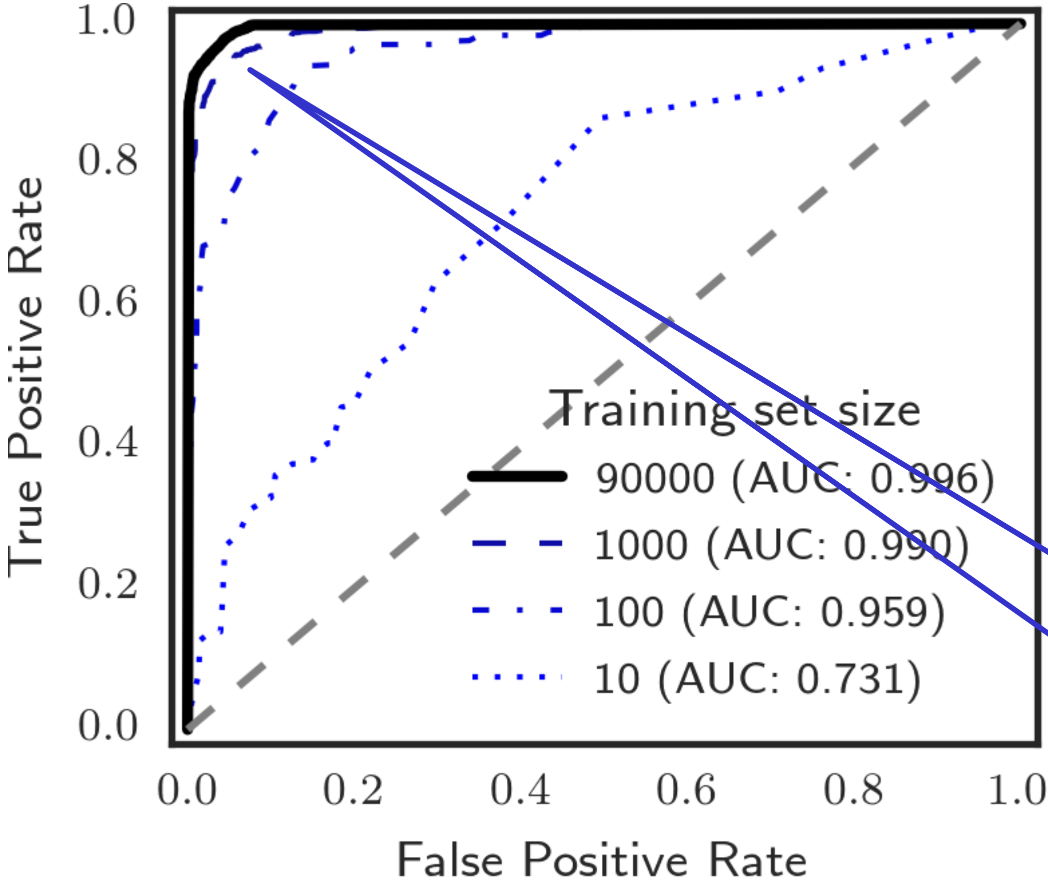


- Flexgrid @ 12.5 GHz
- Transceivers @ 28 GBaud
- 6 Modulation formats
  - (DP) BPSK, QPSK, 8-QAM to 64-QAM,
- Traffic requests: [50;1000] Gbps
- 3 candidate paths per node pair
- BER threshold $T = 4*10^{-3}$

- NB_1: We used synthetic data!
- NB_2: some data sets are becoming available

- *Monia Ghobadi and Ratul Mahajan. "Optical layer failures in a large backbone." In Proceedings of the 2016 Internet Measurement Conference. ACM, 2016.*
- *Rachee Singh, Monia Ghobadi, Klaus-Tycho Foerster, Mark Filer, and Phillipa Gill. "Run, Walk, Crawl: Towards Dynamic Link Capacities." In Proceedings of the 16th ACM Workshop on Hot Topics in Networks. ACM, 2017.*

How long shall training phase be?
(1) *Accuracy vs training set size*



Take-Away 1: Training phase has a reasonable duration

How to reduce required data/probes?

- ML requires training phase with historical data

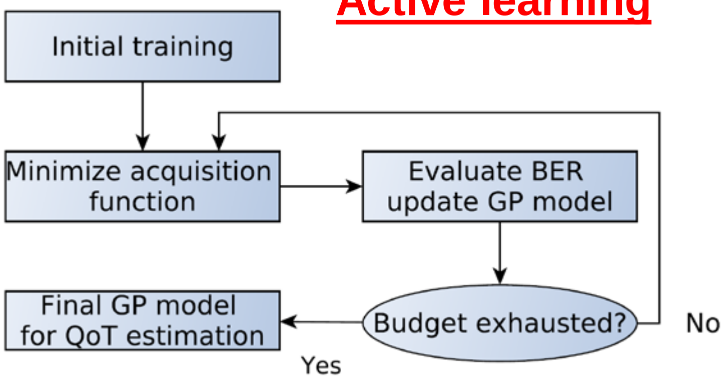
ISSUES

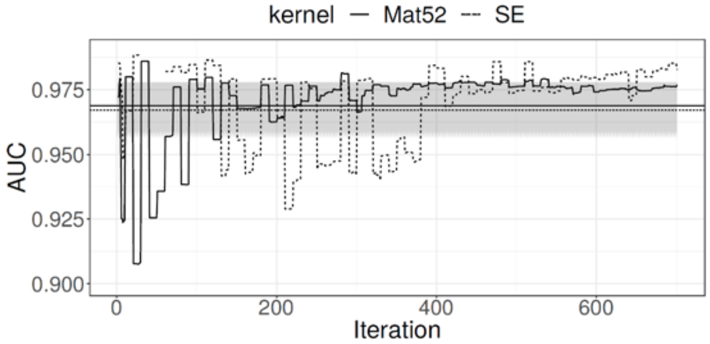
  - Samples from faulty/malfunctioning lightpaths are rare
  - With margined approaches, ligthpaths with risky BER are unlikely deployed (thus never observed)

- Probe ligthtpaths are n… gions of feature space not … data

COSTLY! THE NUMBER OF PROBES MUST BE MINIMIZED!

Active learning

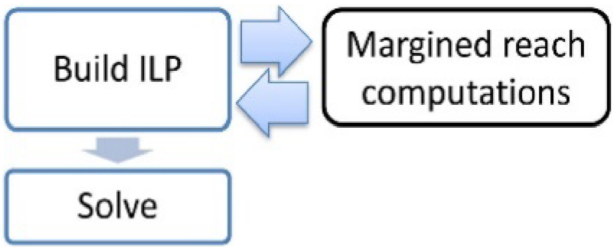


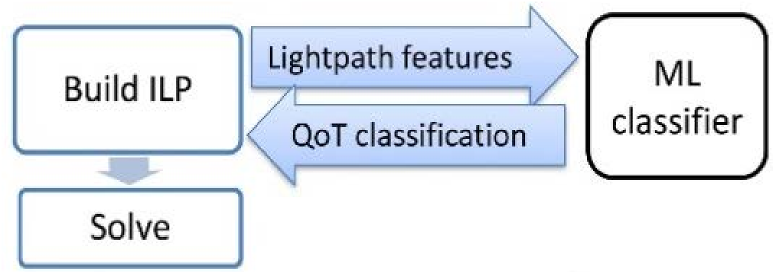


D. Azzimonti, C. Rottondi, M. Tornatore, "Using Active Learning to Decrease Probes for QoT Estimation in Optical Networks," in Proceedings of OFC 2019, San Diego, Feb 2019.

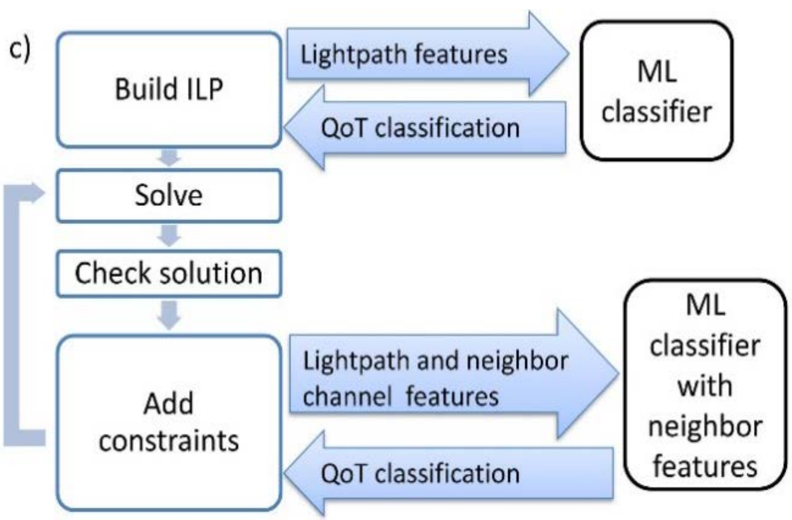
Ok, but, what's the impact on resource saving? Relation between RSA and ML-based QoT estimation

(a) Current mode of operation





(c) Low-margin design (iterative procedure)



Output of the classifier:

probabilty γ that BER ≤ T*

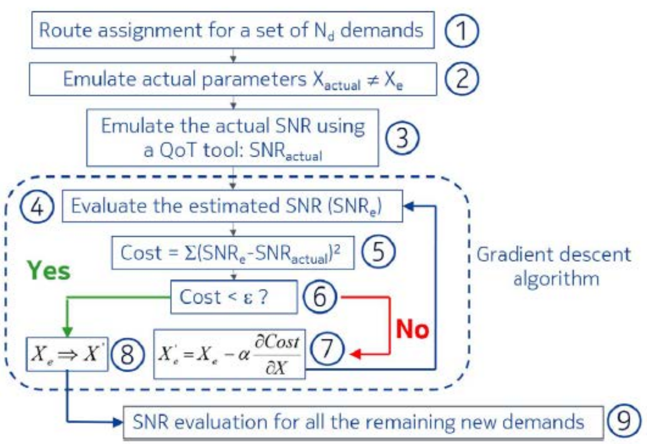
γ: Risk you are willing to accept

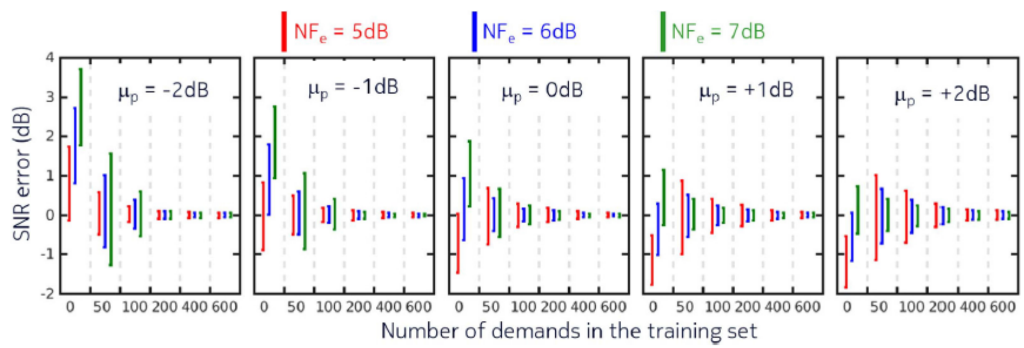
γ	Savings
0.5	35.71%
0.7	32.08%
0.9	27.36%
0.99	26.61%

M. Salani, C. Rottondi, M. Tornatore, "Routing and Spectrum Assignment Integrating Machine-Learning-Based QoT Estimation in Elastic Optical Networks," in Proceedings of INFOCOM 2019, Paris, April 2019.

Another way of looking at this problem..
Estimating unknown parameters in GN model

- General motivation
  - If you have a model, you should use it!
  - No need to reinvent the wheel
- So, for QoT estimation, if we know which parameter is inaccurate (e.g., noise figure), we can use ML to estimate that parameter, and mantain the rest of the analytical model





E. Seve, J. Pesic, C. Delezoide, S. Bigo, and Y. Pointurier, "Learning Process for Reducing Uncertainties on Network Parameters and Design Margins," J. Opt. Commun. Netw. 10, A298-A306 (2018)

Similar concept in: S. Oda, M. Miyabe, S. Yoshida, T. Katagiri, Y. Aoki, T. Hoshida, J. C. Rasmussen, M. Birk, and K. Tse, "A learning iving network with open ROADMs," J. Lightwave Technol., vol. 35, pp. 1350–1356, 2017

Experimental demonstration in multi-domain networks with alien wavelenghts

- QoT estimation is challenging in **multidomain** networks, as each domain administrator discloses very limited intradomain information
- Authors estimate directly OSNR using NNs
  - Note: regression vs classification

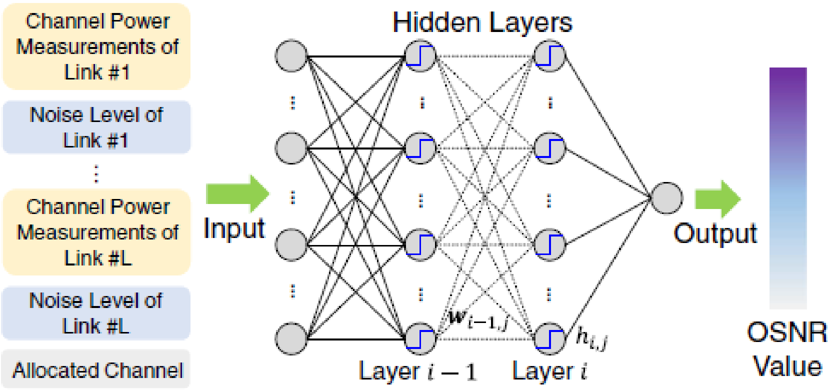


Fig. 7. Structure of the OSNR estimator.

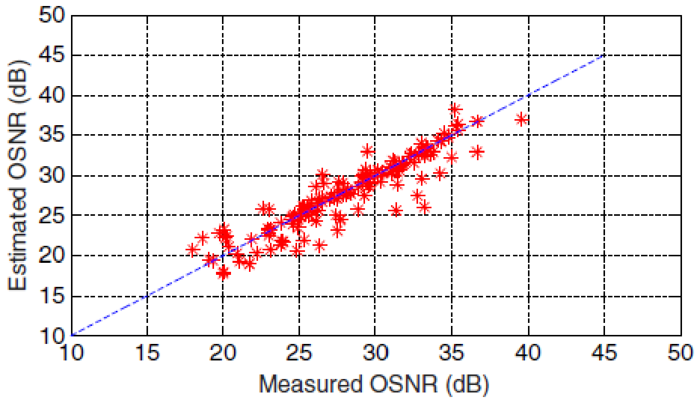


Fig. 9. Comparison between measured (blue dashed line) and estimated (red stars) OSNR.

R. Proietti et al., "Experimental demonstration of machine-learning-aided QoT estimation in multi-domain elastic optical networks with alien wavelengths," in IEEE/OSA J. of Optical Comm. and Netw., vol. 11, no. 1, pp. A1-A10, Jan. 2019.

On dataset dimension for QoT estimation

Different networks, different behaviors!

TABLE IV
NUMBER OF EXAMPLES IN THE EVALUATION SETS, $\overline{S}_p$

p		GBN	TIM	SPARKLE	CORONET
0.1%	Positive	—	2478	556	265
	Negative	—	2840	605	265
	Total	—	5318	1161	530
1%	Positive	105	18,697	3193	1904
	Negative	84	28,359	6030	2635
	Total	189	47,056	9223	4539
5%	Positive	437	72,789	9144	6478
	Negative	412	141,785	30,140	13,165
	Total	849	214,574	39,284	19,643
10%	Positive	745	129,503	15,172	10,243
	Negative	820	283,570	60,280	26,330
	Total	1565	413,073	75,452	36,573
50%	Positive	2848	498,104	27,232	31,307
	Negative	4008	1,206,895	301,395	131,650
	Total	6856	1,704,999	328,627	162,957
100%	Positive	8065	1,272,926	27,253	38,021
	Negative	12,430	4,398,434	1,178,322	488,579
	Total	20,495	5,671,360	1,205,575	526,600

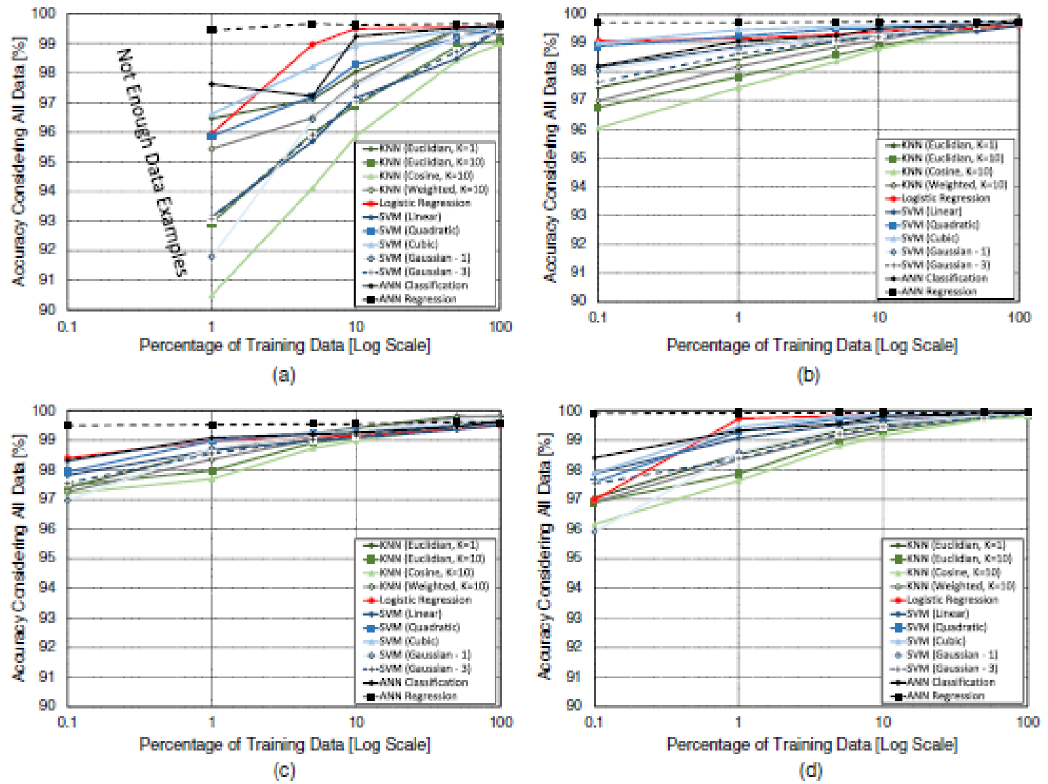


Fig. 7. Accuracy predicting S using machine learning models trained with 0.1%, 1%, 5%, 10%, 50%, and 100% of S for (a) GBN, (b) TIM, (c) SPARKLE, and (d) CORONET.

R. M. Morais and J. Pedro, "Machine learning models for estimating quality of transmission in DWDM networks," in IEEE/OSA Journal of Optical Communications and Networking, vol. 10, no. 10, pp. D84-D99, Oct. 2018

Outline: Part 2
Applications

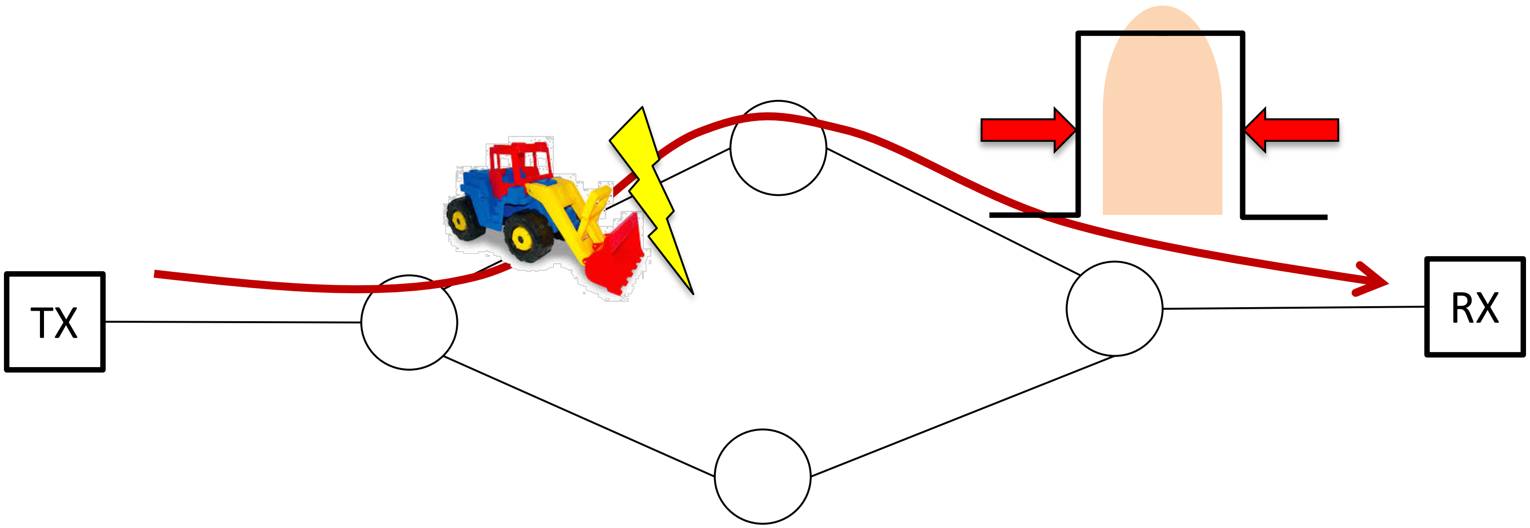
1. ML for QoT Estimation for Unestablished Lighpaths
   - C. Rottondi, L. Barletta, A. Giusti and M. Tornatore, *A Machine Learning Method for QoT Estimation of Unestablished Lightpaths*, in IEEE/OSA Journal of Optical Comm.& Netw. Vol. 10, No. 2, Feb. 2018

2. ML for Failure Management
   - Francesco Musumeci ,et al., “A Tutorial on Machine Learning for Failure Management in Optical Networks”, in IEEE/OSA Journal of Lightwave Technology, available online

3. An overview of other applications at network layer
   - F. Musumeci et al., “A Survey on Application of Machine Learning Techniques in Optical Networks”, Submitted to IEEE communication surveys and tutorials, available in ArXiv
   - Javier Mata, et a., Artificial intelligence (AI) methods in optical networks: A comprehensive survey, Optical Switching and Networking, Volume 28, 2018, pp. 43-57

Two main failure types in optical networks

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

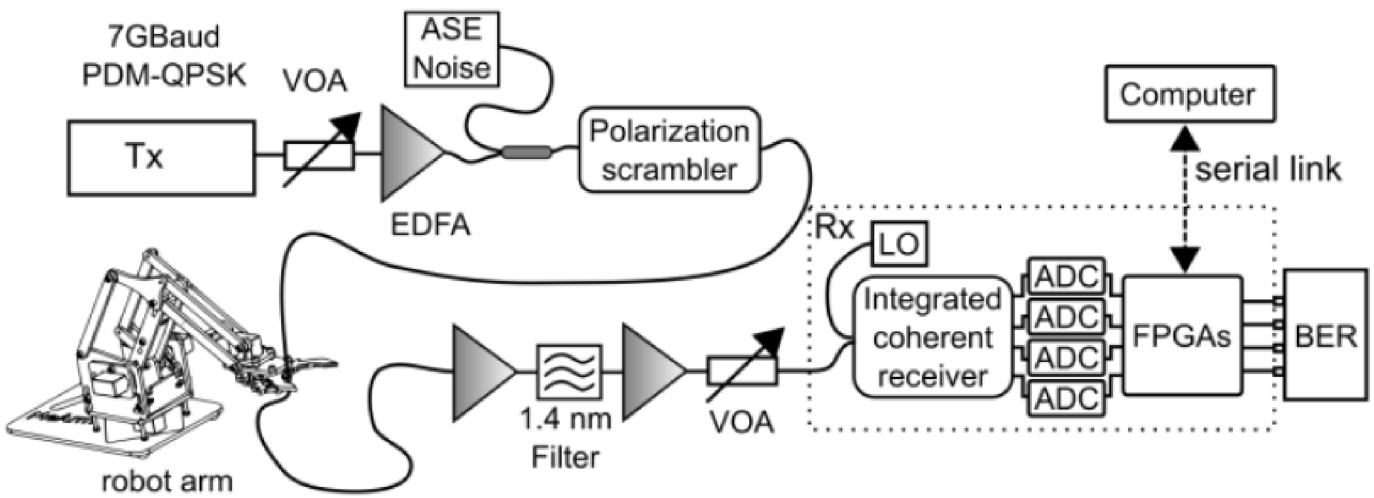


***F. Boitier *et al*., "Proactive Fiber Damage Detection in Real-time Coherent Receiver," *2017 European Conference on Optical Communication (ECOC)*, Gothenburg, 2017, pp. 1-3.**

… Something can be done for hard failures!

Proactive fiber damage detection [1,2,3]

- *"Algorithm extension for a coherent receiver, coupled with machine learning, to monitor mechanical stress of an optical fiber, for recognizing fiber breaks before they occur"*
  - Monitoring of State of Polarization (SOP) of an out-of-band unmodulated laser light
  - Demonstrated 95% accuracy over real-time PDM-QPSK testbed
  - No additional hardware thanks to DSP in receIver



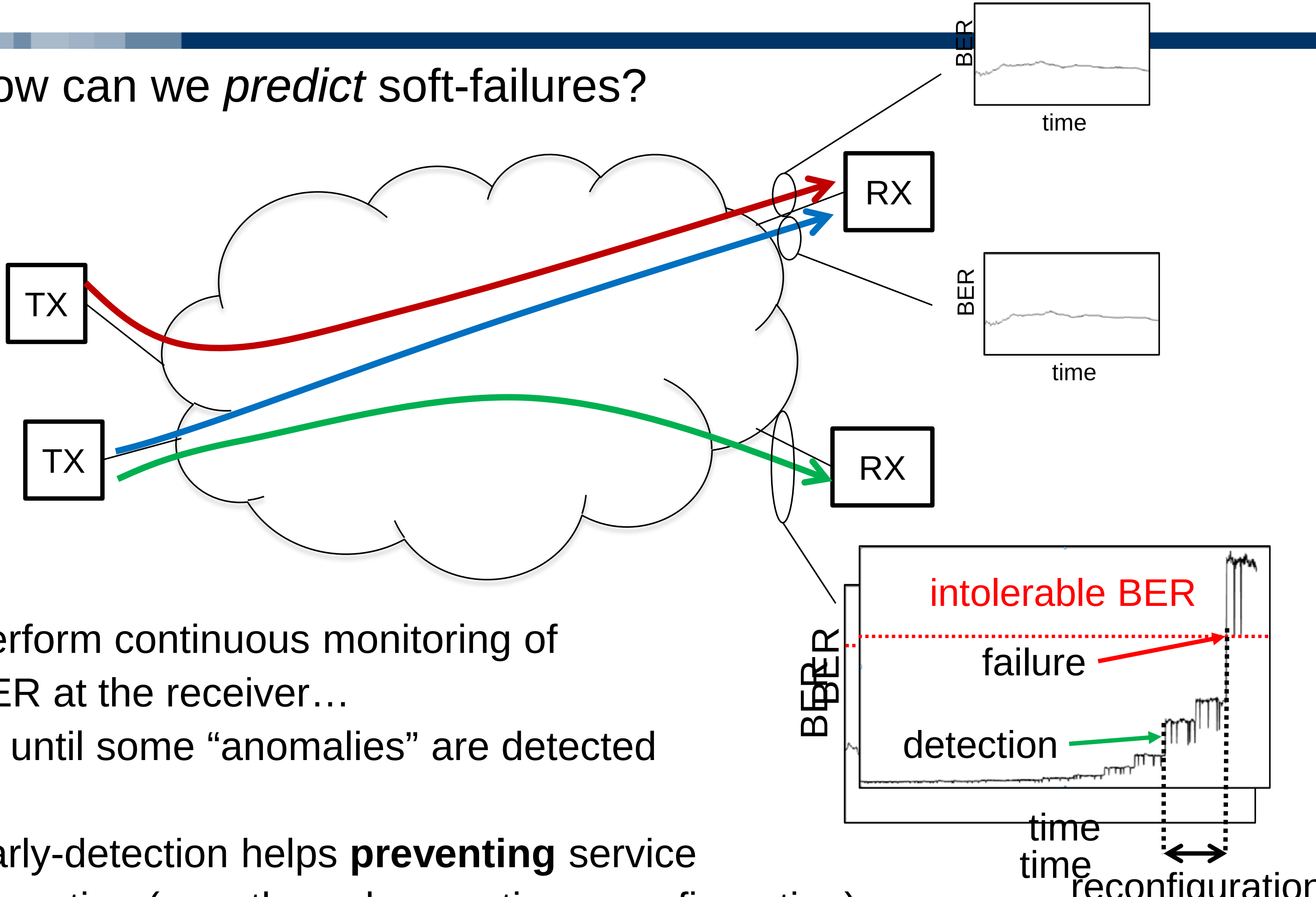
F. Boitier *et al.*, "Proactive Fiber Damage Detection in Real-time Coherent Receiver," *2017 European Conference on Optical Communication (ECOC)*, Gothenburg, 2017, pp. 1-3

J. Pesic et al., "Proactive restoration of optical links based on the classification of events," Proc. ONDM, (2011).

J. E. Simsarian et al., "Shake Before Break: Per-Span Fiber Sensing with In-Line Polarization Monitoring," Proc. OFC, M2E.6 (2017)

Soft-failure *early detection*

- How can we *predict* soft-failures?



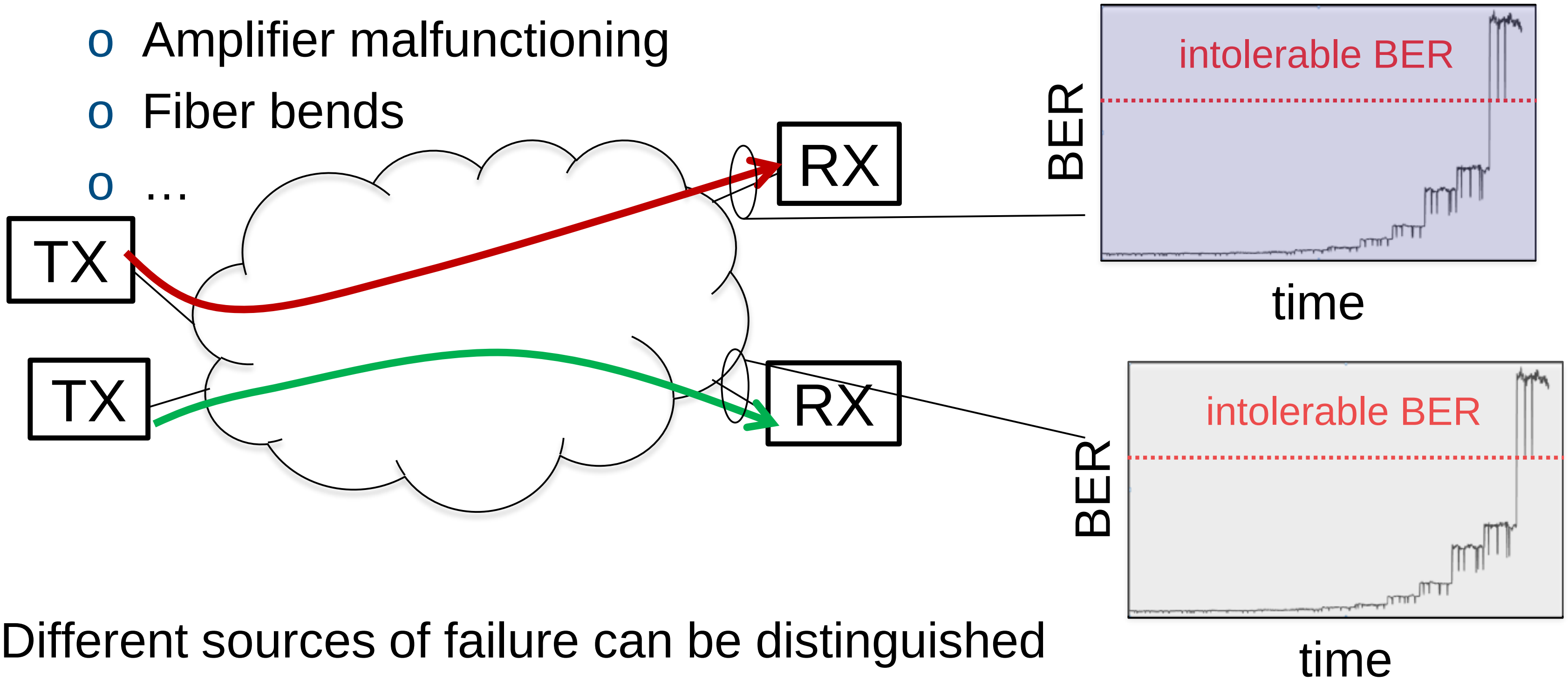
Perform continuous monitoring of BER at the receiver…
… until some "anomalies" are detected

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

A. Vela et al., "BER degradation Detection and Failure Identification in Elastic Optical Networks", in IEEE/OSA Journal of Lightwave Technology, vol. 35, no. 21, pp. 4595-4604, Nov.1, 1 2017.

Soft-failure *cause identification*

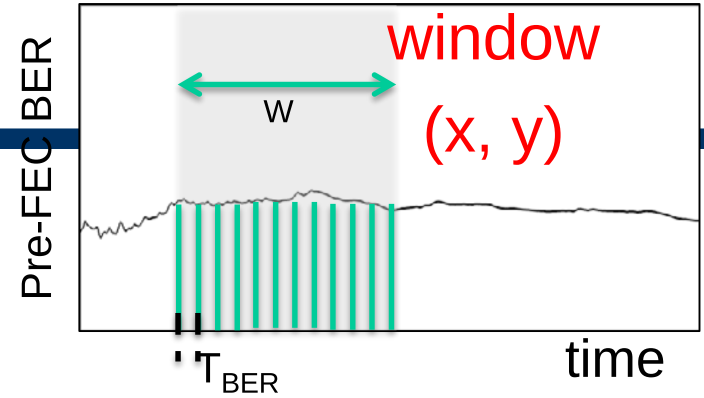
- How can we identify the *cause* of the failure?
  - Failures can be caused by different sources
    - Filters shrinking/misalignment
    - Amplifier malfunctioning
    - Fiber bends
    - …



Different sources of failure can be distinguished via the different effects on BER (i.e., via different BER "features")

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

Main Phase of our study
Tuning ML algorithm



1. Data Retrieval

3 decisions

Validation (optimization of hyperparameters)

BER window → Features → ML algorithm

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

- mean
- min/max
- standard dev.

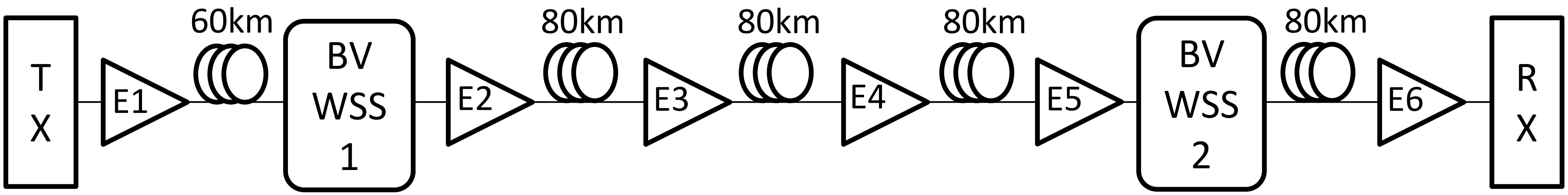
Fault detection:
- Binary SVM
- Random Forest
- Multiclass SVM
- Neural Network

Fault identification:
- Neural Network

3. Prediction and Evaluation

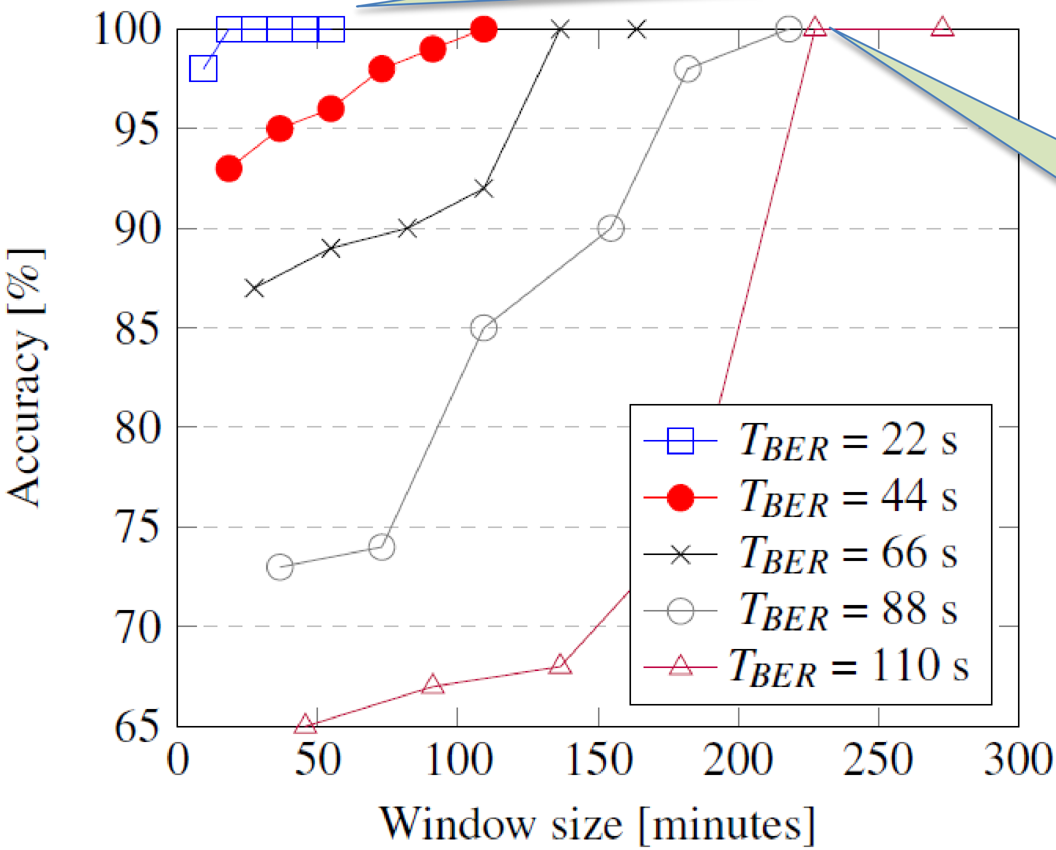
Testbed setup

- 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**)



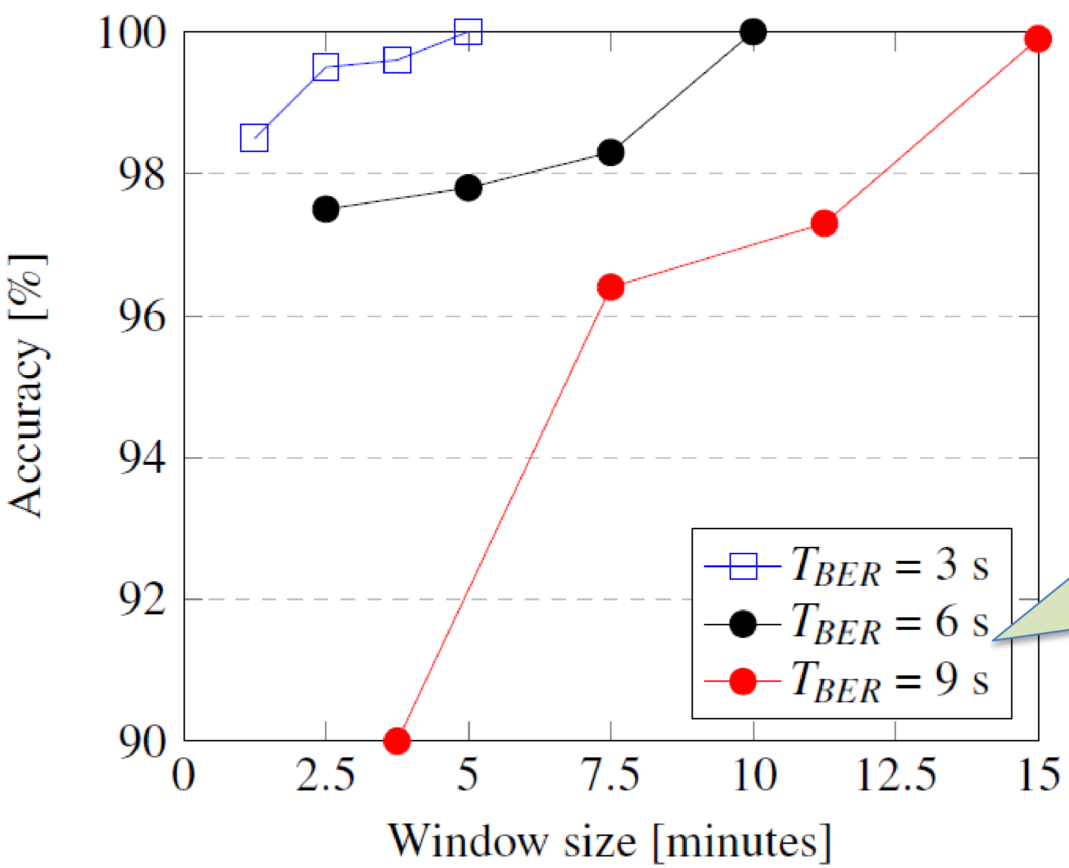
Numerical results: *Detection*
Accuracy vs window features

Take-away 1: Higher performance for with low sampling time
→ Fast monitoring equipment is required



Take-away 2: For increasing sampling time, longer "Windows" are needed for high accuracy

Numerical results: *Identification*
Accuracy vs window features

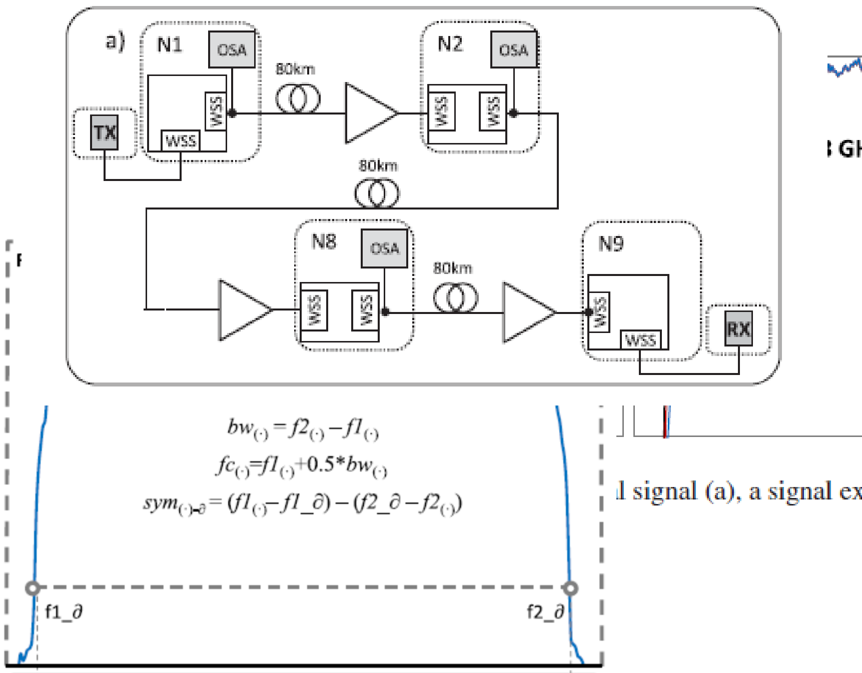


Take-away 3: To perform failure-cause identification, much smaller sampling period is needed wrt failure detection

Detection and Identification using Optical Spectrum Analyzers

- Cost-effective Optical Spectrum Analyzers (OSA) with sub-GHz resolution can be used to monitor spectrum along transmission line

Flexgrid High Resolution Optical Channel Monitor (OCM). [Online]. Available: http://www.finisar.com, Accessed: Jun. 2018.



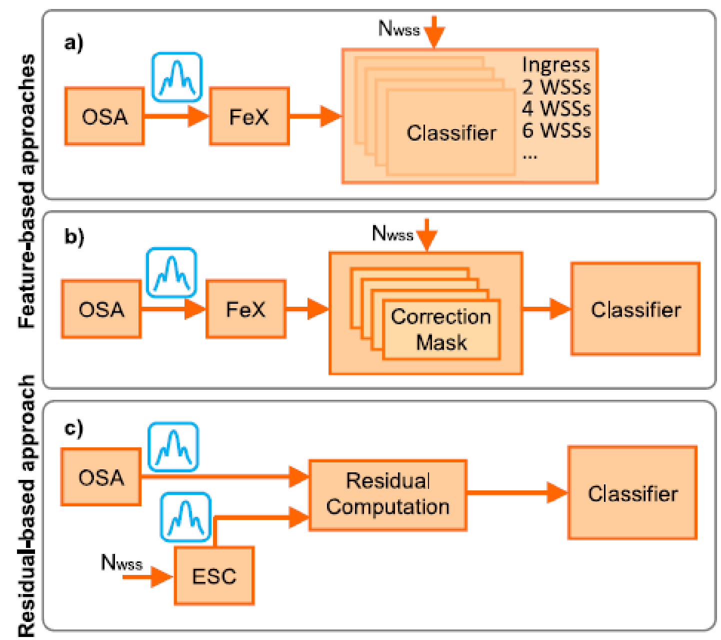


Fig. 4. Approaches to solve the filter cascading problem: (a) multi-classifier, (b) single-classifier, and (c) residual computation.

B. Shariati, M. Ruiz, J. Comellas and L. Velasco, "Learning From the Optical Spectrum: Failure Detection and Identification," in Journal of Lightwave Technology, vol. 37, no. 2, pp. 433-440, 15 Jan.15, 2019.

Detection of attacks (I)

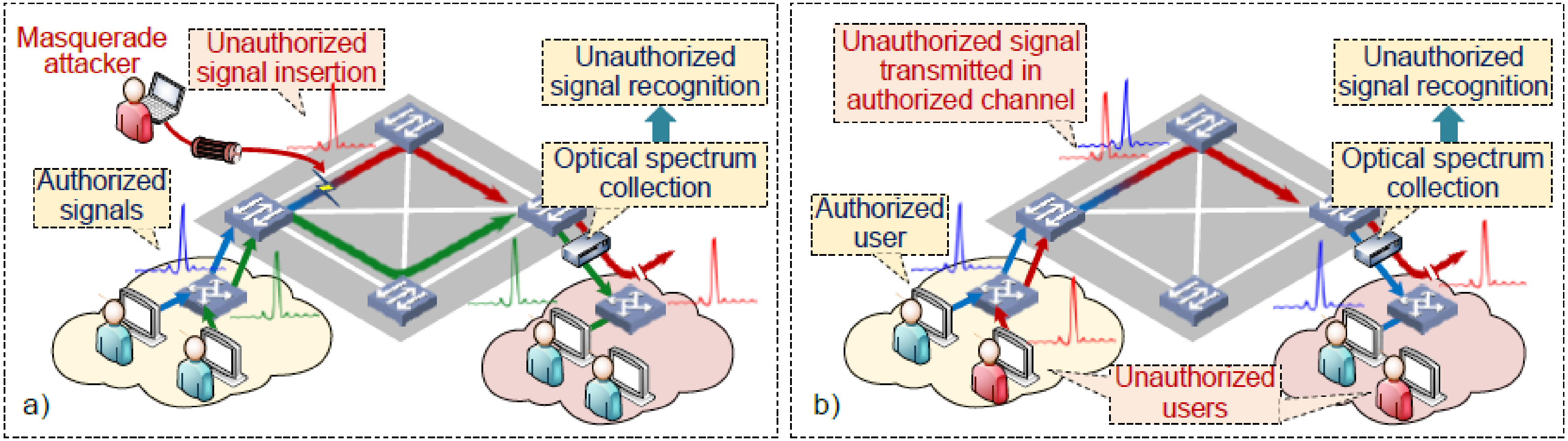


Fig. 1. (a) Masquerade attacker gain access to network incognito and insert signals; (b). Unauthorized users transmit unauthorized signals in authorized channels

Y. Li, et al. , Optical spectrum feature analysis and recognition for optical network security with machine learning, Optics Express, to appear

Detection of attacks (II)

- Different (simpler, but less controversial) intuition:
  - Jamming attack affects physical properties legitimate signals
- Feature are typical parameters of coherent receivers:
  - chromatic dispersion (CD)
  - differential group delay (DGD)
  - Optical Signal-to-Noise Ratio (OSNR),
  - Polarization dependent loss (PDL),
  - Q-factor
  - pre-FEC bit errors (BE-FEC),
  - pre-FEC bit error rate (BER-FEC)
  - uncorrected block errors (UBE-FEC)
  - optical power received (OPR)
  - optical power transmitted (OPT)

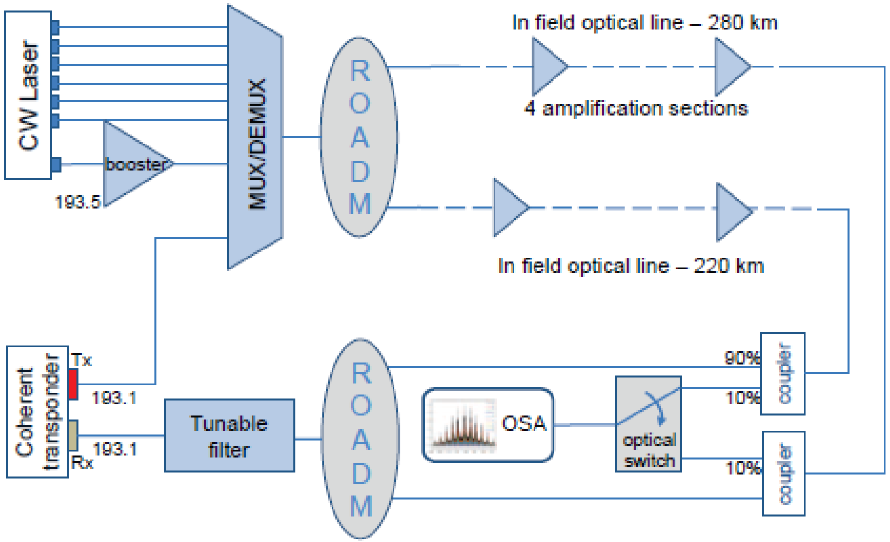


Fig. 2: Setup used in the experiments.

- SVM and ANN reach 100% accuracy [1]
- In case of unknown attacks:
  - Unsupervised learning (Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [2])

[1] C. Natalino, et al. "Field demonstration of machine-learning-aided detection and identification of jamming attacks in optical networks," ECOC, 2018

[2] M. Furdek, et al. "Experiment-based detection of service disruption attacks in optical networks using data analytics and unsupervised learning." Photonics West, 2019.

Outline: Part 2 Applications

1. ML for QoT Estimation for Unestablished Lighpaths

- C. Rottondi, L. Barletta, A. Giusti and M. Tornatore, *A Machine Learning Method for QoT Estimation of Unestablished Lightpaths*, in IEEE/OSA Journal of Optical Comm.& Netw. Vol. 10, No. 2, Feb. 2018
- D. Azzimonti, C. Rottondi, M. Tornatore, "Using Active Learning to Decrease Probes for QoT Estimation in Optical Networks," in Proceedings of OFC 2019, San Diego, Feb 2019.
- M. Salani, C. Rottondi, M. Tornatore, "Routing and Spectrum Assignment Integrating Machine-Learning-Based QoT Estimation in Elastic Optical Networks," in Proceedings of INFOCOM 2019, Paris, April 2019.

2. ML for Failure Management

- 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
- A. Vela et al., "BER degradation Detection and Failure Identification in Elastic Optical Networks", in IEEE/OSA Journal of Lightwave Technology, vol. 35, no. 21, pp. 4595-4604, Nov.1, 1 2017
- Francesco Musumeci ,et al., "A Tutorial on Machine Learning for Failure Management in Optical Networks", in IEEE/OSA Journal of Lightwave Technology, available online

3. An overview of other applications at network layer

- F. Musumeci et al., "A Survey on Application of Machine Learning Techniques in Optical Networks", Submitted to IEEE communication surveys and tutorials, available in ArXiv
- Javier Mata, et a., Artificial intelligence (AI) methods in optical networks: A comprehensive survey, Optical Switching and Networking, Volume 28, 2018, pp. 43-57

Overview of other applications

Physical layer

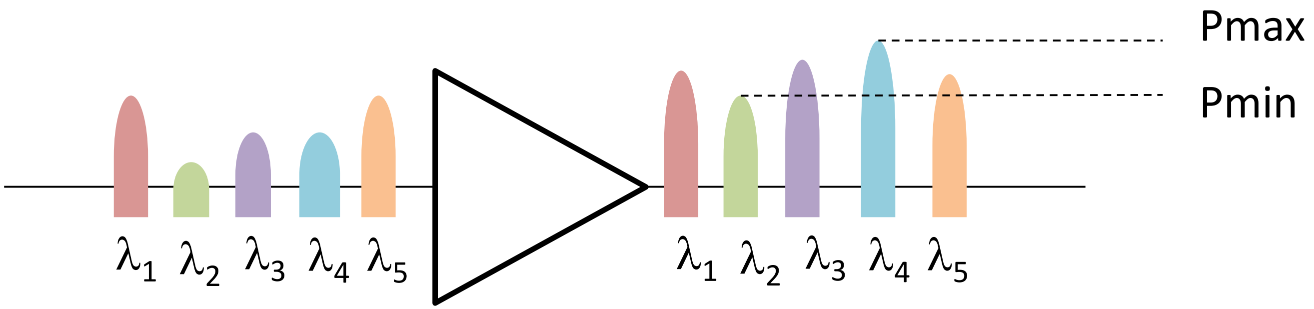
1. Quality of Transmission (QoT) estimation
2. Optical amplifier control
3. Modulation format recognition
4. Nonlinearities mitigation
5. «Sensing»

Network layer

1. Traffic prediction and virtual topology design
2. Failure detection and localization
3. Flow classification

Classification taken from: F. Musumeci et al., "A Survey on Application of Machine Learning Techniques in Optical Networks", IEEE Communication Surveys and Tutorials, 2019

Physical layer
Optical amplifier control

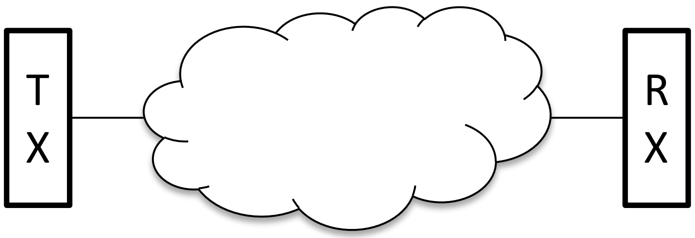


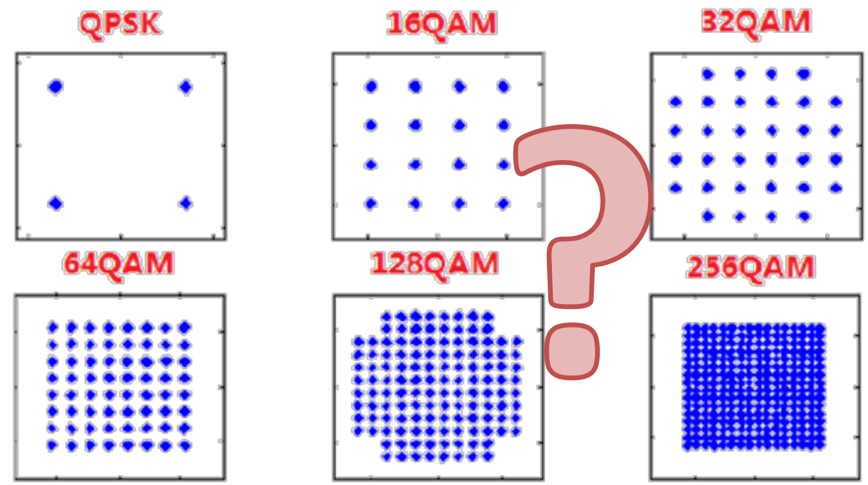
– typically not generalizable
– depend on the specific system (gain-control mechanism, EDFA gain tilt, nr of EDFAs…) which use to vary during their activity

Huang et al., "Dynamic mitigation of EDFA power excursions with machine learning", Optics Express, vol. 25 n. 3, Feb. 2017
Bastos et al., "Mapping EDFA Noise Figure and Gain Flatness Over the Power Mask Using Neural Networks", Journal of Microwaves, Optoelectronics and Electromagnetic Applications, vol. 12, n. SI-2, July 2013

Physical layer

Modulation format recognition (MFR)





– additional delay for in signal detection

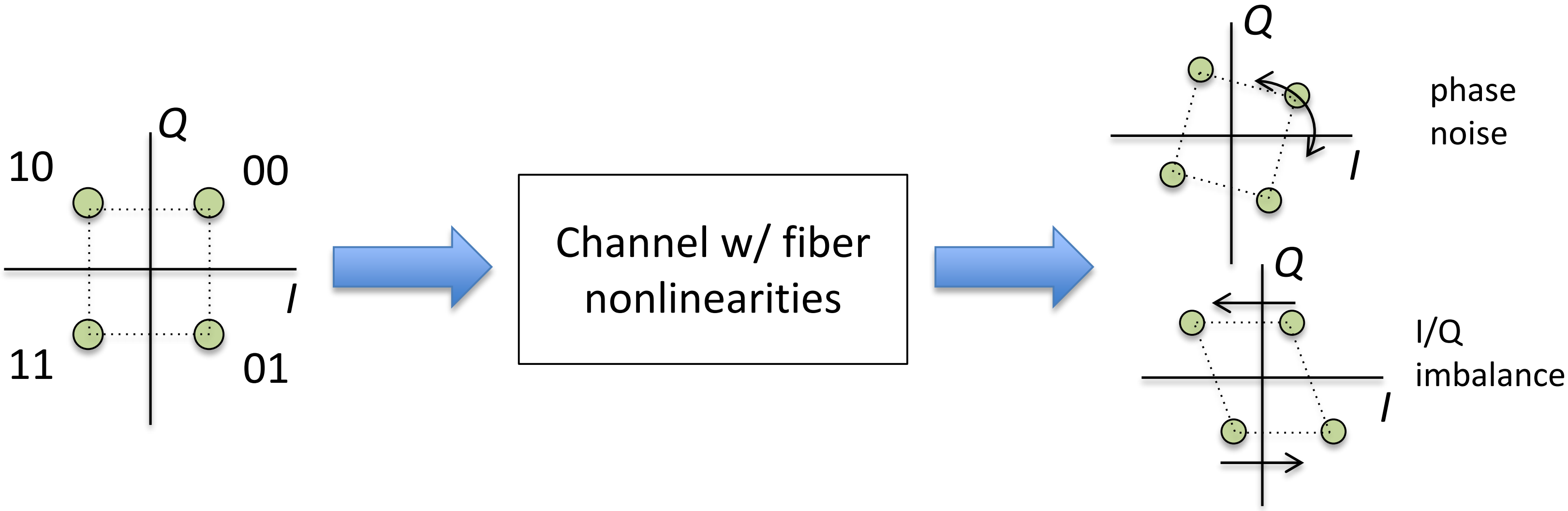
Khan et al., “Modulation Format Identification in Coherent Receivers Using Deep Machine Learning”, Photonics Technology Letters, vol. 28 n. 17, Sep. 2016

Khan et al., “Non-data-aided joint bit-rate and modulation format identification for next-generation heterogeneous optical networks”, Optical Fiber Technology, vol. 20 n. 2, Mar. 2014

Tan et al., “Simultaneous Optical Performance Monitoring and Modulation Format/Bit-Rate Identification Using Principal Component Analysis”, Journal of Optical Communications and Networking, vol. 6 n. 5, May 2014

Physical layer
Nonlinearities mitigation

– Kerr effect, self-phase modulation (SPM), cross-phase modulation (XPM)…



- Traditional methods require complex mathematical models and prior information on the traversed channel
- ML enables “safer” decision by learning from actual channel properties

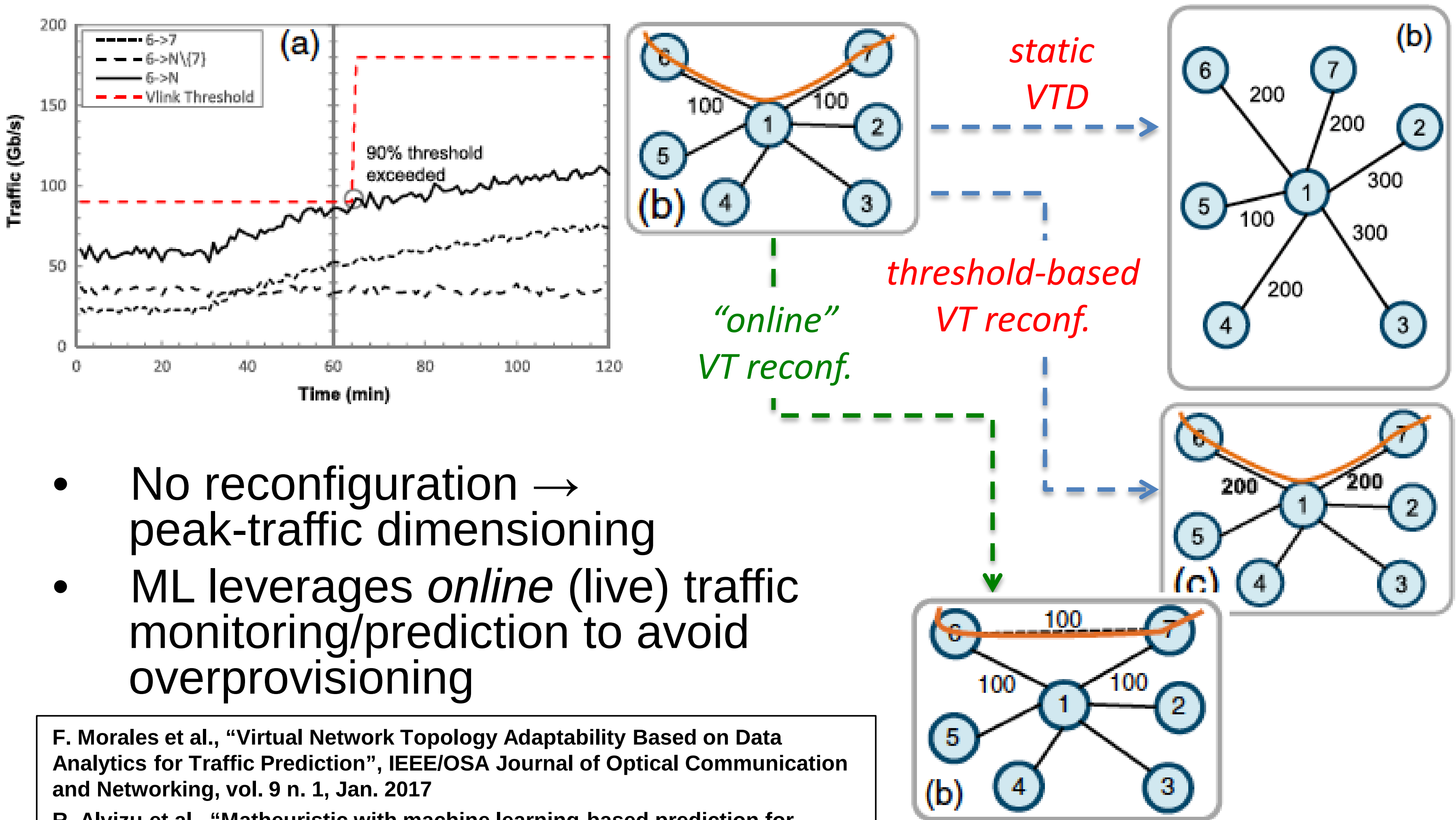
Wang et al., “Nonlinear Decision Boundary Created by a ML-based Classifier to Mitigate Nonlinear Phase Noise”, in ECOC 2015

Wang et al., “Nonlinearity Mitigation Using a ML Detector Based on k-Nearest Neighbors”, Photonics Tech. Letters, 2016

S. Zhang, et. al, “Field and lab experimental demonstration of nonlinear impairment compensation using neural networks,” Nature Communications, 2019

F. Ye, et al., "A new and simple method for crosstalk estimation in homogeneous trench-assisted multi-core fibers," in Asia Communications and Photonics Conference 2014

D. Zibar, et al."Application of machine learning techniques for amplitude and phase noise characterization," J. Lightwave Technol. 33(7), 1333–1343 (2015).



- No reconfiguration → peak-traffic dimensioning
- ML leverages *online* (live) traffic monitoring/prediction to avoid overprovisioning

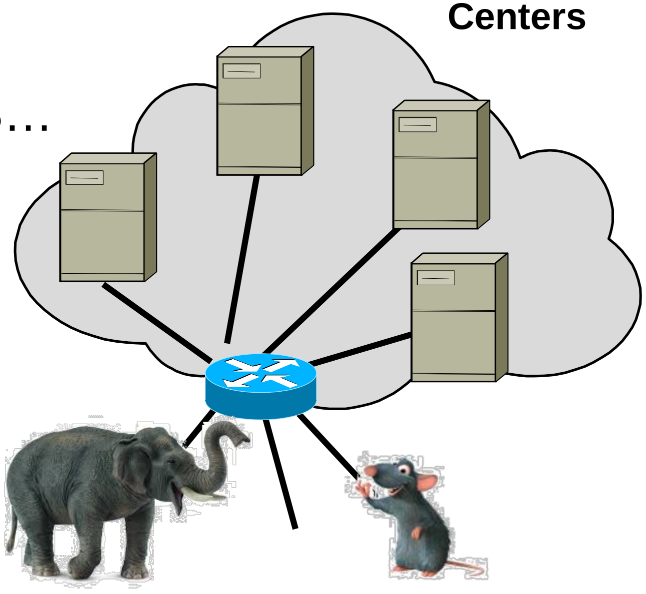
F. Morales et al., "Virtual Network Topology Adaptability Based on Data Analytics for Traffic Prediction", IEEE/OSA Journal of Optical Communication and Networking, vol. 9 n. 1, Jan. 2017

R. Alvizu et al., "Matheuristic with machine learning-based prediction for software-defined mobile metro-core networks", IEEE/OSA Journal of Optical Communication and Networking, vol. 9 n. 9, Sep. 2017

Network layer
Flow classification

- protocols (http, ftp, smtp…)
- services (VoD, data transfer, text messages…)
- requirements (latency, bandwidth, jitter…)
- network “customers” (human end-users, companies, sensors)
  - E.g., “mice” vs “elephant” flows in Data Centers

SLAs, QoS…



L. Wang, X. Wang, M. Tornatore, K. Joon Kim, S.-M. Kim, D.-U Kim, K.-E. Han, and B. Mukherjee, “Scheduling With Machine-Learning-Based Flow Detection for Packet-Switched Optical Datacenter Networks, JOCN2018

Viljoen et al., “Machine Learning Based Adaptive Flow Classification for Optically Interconnected Data Centers”, in ICTON 2016, July 2016

Cao et al., “An accurate traffic classification model based on support vector machines”, International Journal on Network Management, 27:e1962, 2017.

Sensing

An experiment with road traffic

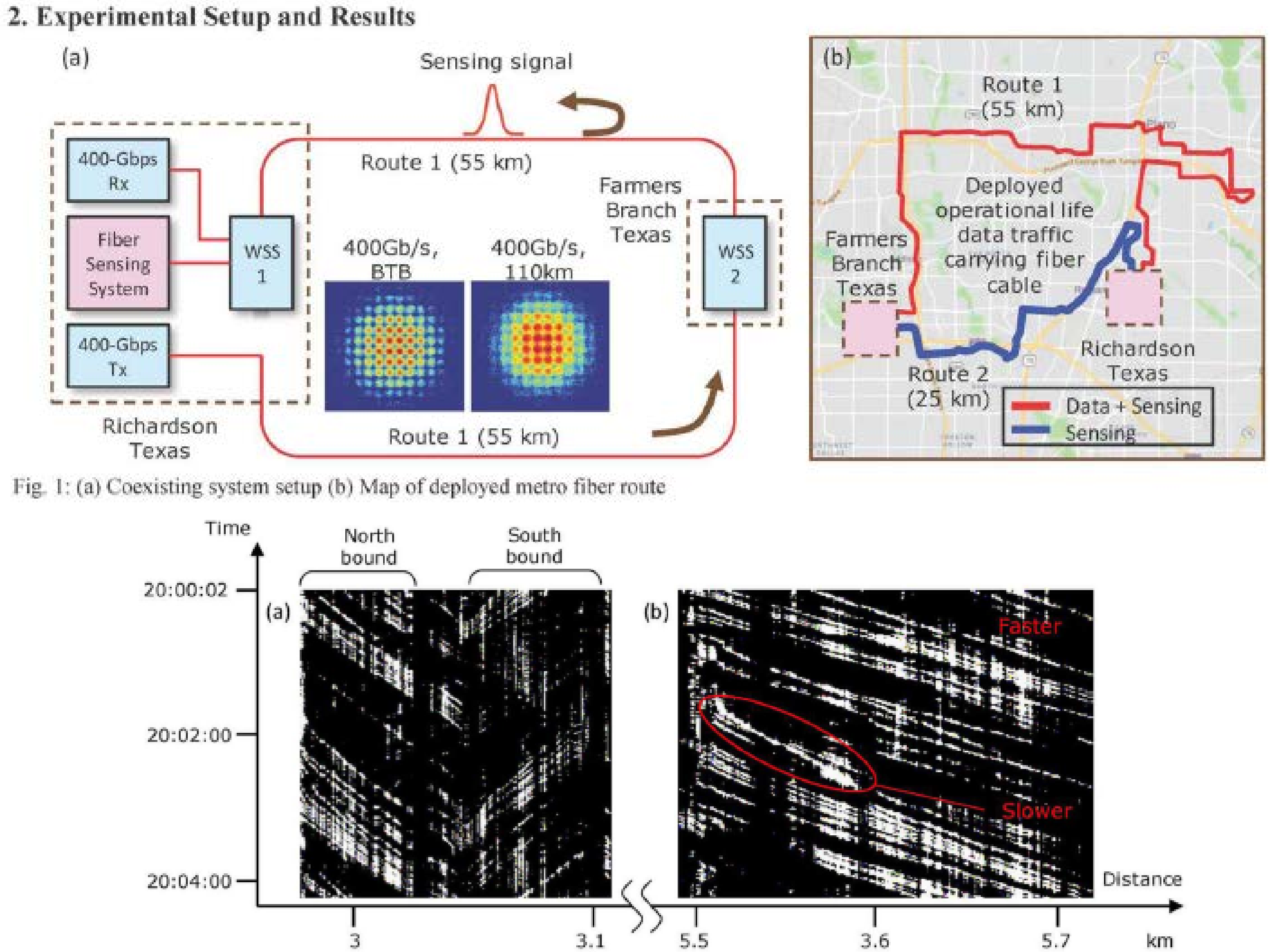


Fig. 1: (a) Coexisting system setup (b) Map of deployed metro fiber route

Fig. 2: Examples of water-fall trace for fiber sensing system.

G. Wellbroock, First Field Trial of Sensing Vehicle Speed, Density, and Road Conditions by using Fiber Carrying High Speed Data, post-deadline, OFC 2019

These slides are not NIST's slides. NIST is not responsible for the content of these slides.

Conclusion: My personal opinion

- QoT estimation
  - Partly. Ok for improving accuracy, or when unknowns are too many
- Failure management
  - Yes! Root cause analysis (it is a complex semisupervised problem!)
- Traffic prediction
  - Yes! (Check DC-NN*)
- Resource allocation (e.g., dynamic traffic allocation)
  - Skeptical
    - Several problems (traffic varies, scalability…)
- Sensing

*D. Andreoletti, S. Troia, F. Musumeci, S. Giordano, G. Maier, M. Tornatore, «Network Traffic Prediction based on Diffusion Convolutional Recurrent Neural Networks», Infocom 2019

Some emerging concepts (I)

- **Active Learning**
  - *No explicit separation between training and testing*, continuos training as new data arrives
  - Great in situation where data is scarce expensive

- P. Deisenroth, D. Fox, and C. E. Rasmussen, “Gaussian processesfor data-efficient learning in robotics and control,”IEEE Transactions onPattern Analysis and Machine Intelligence, vol. 37, no. 2, pp. 408–423,Feb 2015
- D. Azzimonti, C. Rottondi, and M. Tornatore, “Using Active Learning to Decrease Probes for QoT Estimation in Optical Networks,” in Optica lFiber Communications Conference (OFC), Mar. 2019

- **Transfer Learning**
  - *Is the training performed over a network/link/failure still valid on a different scenario?*

- ftp://ftp.cs.wisc.edu/machine-learning/shavlik-group/torrey.handbook09.pdf
- W Mo, YK Huang, S Zhang, E Ip, DC Kilper, Y Aono, T Tajima, «ANN-based transfer learning for QoT prediction in real-time mixed line-rate systems», in Optical Fiber Communications Conference (OFC), Mar. 2019

Some emerging concepts (II)

- **Interpretability**
  - Go beyond black-box machine learning outcome!
  - Can we gain insights on our problems

- H.J. Escalante, I. Guyon, S. Escalera X. Baro, Y. Gucluturk, U. Guclu and M. van Gerven, Explainable and Interpretable Models in Computer Vision and Machine Learning, *Springer Series on Challenges in Machine Learning*, 2018.
- F. N. Khan, Q. Fan, C. Lu and A. P. T. Lau, "An Optical Communication's Perspective on Machine Learning and Its Applications," in Journal of Lightwave Technology, vol. 37, no. 2, pp. 493-516, 15 Jan.15, 2019.

- **Collaborative Self-Learning**
  - Different network nodes perform local estimations, then share part of their local knowledege to improve overall knowledge of other nodes
  - 4 phases: i) knowledge discover; ii) knowledge share; iii) knowledge assimilate; and iv) knowledge usage

- M. Ruiz, F. Boitier, P. Layec, and L. Velasco, "Self-Learning Approaches for Real Optical Networks," in Proc. IEEE/OSA Optical Fiber Communication Conference (OFC), 2019

Thank You!

POLITECNICO MILANO 1863

Achille Pattavina
Francesco Musumeci
Shahin Shahkarami
Luca Barletta

THE UNIVERSITY OF CALIFORNIA DAVIS

Biswanath Mukherjee, Yu Wu
Lin Wang,Sabidur Rehman,

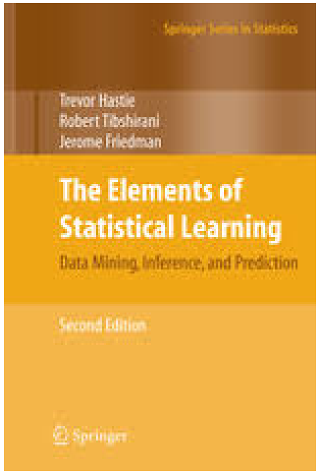
Filppo Cugini (CNIT)
Cristina Rottondi (PoliTo)
Dario Azzimonti, Matteo Salani, Alessandro Giusti
(Dalle Molle Institute of Artificial Intelligence)

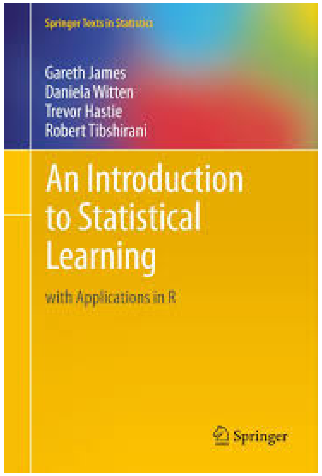
H2020 Metro Haul

European Commission | Horizon 2020 European Union funding for Research & Innovation

Some material

- Books (general refs. for ML):
  - T. Hastie, R. Tibshirani, J. Friedman, "The Elements of Statistical Learning", Ed. Springer
  - G. James, D. Witten, T. Hastie, R. Tibshirani, "An Introduction to Statistical Learning with Applications in R", Ed. Springer





- Prof. Andrew Ng lectures (Stanford University)
- … Google it!

Some publications (1)

e Miguel *et al.*, “Cognitive Dynamic Optical Networks”, *Journal of Optical Communication and Networking*, vol. 5, n. 10, Oct. 2013

et al., “Machine Learning Techniques for Optical Performance Monitoring From Directly Detected PDM-QAM Signals”, *Journal of Lightwave Technology*, vol. 35, n. 4, Feb. 2017

et al., “Experimental demonstration of a cognitive quality of transmission estimator for optical communication systems”, *Optics Express*, vol. 20, n. 26, Dec. 2012

et al., “A Cognitive Quality of Transmission Estimator for Core Optical Networks”, *Journal of Lightwave Technology*, vol. 31, n. 6, Mar. 2013

Some publications (2)