Technische Universität München

TUM School of Computation, Information and Technology

Planning and Control of Disaggregated Quantum-Safe Optical Transport Networks

Sai Kireet Patri, M.Sc.

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften (Dr.-Ing.)

genehmigten Dissertation.

Vorsitz:		Prof. Dr.-Ing. Norbert Hanik
Prüfer*innen der Dissertation:	1.	Prof. Dr.-Ing. habil. Carmen Mas Machuca
	2.	Prof. Marija Furdek Prekratic, Ph.D.

Die Dissertation wurde am 13.04.2023 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 04.08.2023 angenommen.

Planning and Control of Disaggregated Quantum-Safe Optical Transport Networks

Sai Kireet Patri, M.Sc.

04.08.2023

Abstract

Optical communication networks have become an essential element of long-distance and high-speed communication since their widespread deployment in the last two decades. The migration towards cloud data centers has increased the need for high-speed, high-capacity interconnection between critical sites. Moreover, as fifth-generation (5G) networks are being rapidly deployed across the globe and sixth-generation (6G) network requirements are currently being defined, network operators and service providers need to profitably deploy high-capacity optical networking devices.

However, traditional optical networks based on a fixed-grid spectrum do not scale proportionally to such growth in demand and require expensive upgrades in terms of optical components and transmission bands. This dissertation introduces and evaluates heuristics and combinatorial optimization algorithms to plan optical networks capable of carrying high-capacity internet traffic over large geographical distances. Network studies conducted across different networks provide pivotal insight into the benefits and drawbacks of various upgrade decisions available to a long-haul network operator. Apart from studying the effects of different flexible optical terminals, the study of optical band transmission upgrades is also conducted for multiple reference optical networks. These studies show a clear advantage of using a combination of multiple optical transmission bands, along with flexible bandwidth variable transceivers, for meeting the traffic demands over the entire planning period.

With the increased deployment of software-tunable optical components, the risk of device failure also increases. To mitigate this risk, a monitoring and fault localization mechanism is explored in this dissertation. The concept of network monitoring and device flexibility can be extended to include device disaggregation, where different optical device vendors adhere to open-source models, thereby driving down network deployment costs. Such deployments need to be emulated for network-wide effects, and cross-device functionality can be ironed out. In this dissertation, the control plane and orchestration of an open-source disaggregated optical device are deployed as a simulation, and several operations are tested.

To guarantee their users' protection against data theft and espionage, network operators are exploring the deployment of quantum-safe approaches in their flexible optical networks.

The devices needed for such a network should also be planned, and the keys generated by such quantum-safe techniques need to be managed and controlled using existing network management principles. In this dissertation, we first plan the device placement for such a quantum-safe network and then delve into key management methodologies to enable end-to-end security against quantum computer attacks.

Kurzfassung

Optische Kommunikationsnetze sind seit ihrer weit verbreiteten Einführung in den letzten zwei Jahrzehnten zu einem wesentlichen Element der Langstrecken- und Hochgeschwindigkeitskommunikation geworden. Der Übergang zu Cloud-Rechenzentren hat den Bedarf an Hochgeschwindigkeitsverbindungen mit hoher Kapazität zwischen kritischen Standorten erhöht. Da die Netze der fünften Generation (5G) weltweit rasch eingeführt werden und die Anforderungen an die Netze der sechsten Generation (6G) derzeit definiert werden, müssen Netzbetreiber und Dienstanbieter optische Netzwerkgeräte mit hoher Kapazität rentabel einsetzen.

Herkömmliche optische Netze, die auf einem Spektrum mit festem Raster basieren, lassen sich jedoch nicht proportional zu einer solchen Nachfragesteigerung skalieren und erfordern teure Aufrüstungen in Bezug auf optische Komponenten und Übertragungsbänder. In dieser Dissertation werden Heuristiken und kombinatorische Optimierungsalgorithmen zur Planung optischer Netze vorgestellt und bewertet, die in der Lage sind, Internetverkehr mit hoher Kapazität über große geografische Entfernungen zu übertragen. Netzstudien, die in verschiedenen Netzen durchgeführt wurden, geben einen entscheidenden Einblick in die Vor- und Nachteile verschiedener Aufrüstungsentscheidungen, die einem Betreiber eines Langstreckennetzes zur Verfügung stehen. Neben der Untersuchung der Auswirkungen verschiedener flexibler optischer Endgeräte wird auch die Aufrüstung des optischen Bandes für mehrere optische Referenznetze untersucht. Diese Studien zeigen einen klaren Vorteil der Verwendung einer Kombination aus mehreren optischen Übertragungsbändern zusammen mit flexiblen, bandbreitenvariablen Transceivern, um die Verkehrsanforderungen über den gesamten Planungszeitraum zu erfüllen.

Mit dem zunehmenden Einsatz von softwareabstimmbaren optischen Komponenten steigt auch das Risiko eines Geräteausfalls. Um dieses Risiko zu mindern, wird in dieser Dissertation ein Mechanismus zur Überwachung und Fehlerlokalisierung erforscht. Das Konzept der Netzüberwachung und der Geräteflexibilität kann auf die Disaggregation von Geräten ausgedehnt werden, wobei verschiedene Hersteller optischer Geräte sich an Open-Source-Modelle halten und so die Kosten für die Netzeinführung senken. Solche Implementierungen müs-

sen emuliert werden, um netzweite Auswirkungen zu erzielen, und die geräteübergreifende Funktionalität kann verbessert werden. In dieser Dissertation werden die Steuerebene und die Orchestrierung eines disaggregierten optischen Open-Source-Geräts als Simulation eingesetzt und verschiedene Operationen getestet.

Um den Schutz ihrer Nutzer vor Datendiebstahl und Spionage zu gewährleisten, untersuchen Netzbetreiber den Einsatz von quantensicheren Ansätzen in ihren flexiblen optischen Netzen. Die für ein solches Netz benötigten Geräte müssen ebenfalls geplant werden, und die mit solchen quantensicheren Techniken erzeugten Schlüssel müssen mit den bestehenden Netzverwaltungsprinzipien verwaltet und kontrolliert werden. In dieser Dissertation planen wir zunächst die Geräteplatzierung für ein solches quantensicheres Netz und befassen uns dann mit Methoden der Schlüsselverwaltung, um eine durchgängige Sicherheit gegen Quantencomputerangriffe zu ermöglichen.

Contents

Acronyms

ACF-EGN Accurate Closed Form Enhanced Gaussian Noise Model 46, 47, 48, 50, 51, 54, 138

ANN Artificial Neural Network 99, 103, 104, 105, 106, 107, 108, 143, 171

APD Avalanche Photo Diode 42

API Application Programming Interface 10, 11, 13, 18, 37, 38, 39, 40, 98, 109, 113, 116, 125, 127, 128, 135, 136, 171

ARIMA Autoregressive Integrated Moving Average 144, 145

ARIMAX Autoregressive Integrated Moving Average with Explanatory Variable 99, 103, 105, 106, 108, 143, 145, 171

ASE Amplified Spontaneous Emission 8, 24, 25, 59, 137

BDM Band Division Multiplexing 68, 70, 73

BER Bit Error Rate 98, 101, 102, 103, 107

BVT Bandwidth Variable Transceiver 1, 2, 7, 8, 9, 10, 12, 17, 18, 19, 21, 23, 25, 26, 27, 28, 29, 39, 45, 46, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 65, 66, 67, 68, 70, 71, 72, 73, 74, 76, 77, 78, 79, 80, 81, 82, 83, 84, 89, 92, 95, 97, 98, 100, 101, 133, 134, 135, 136, 167, 168, 171

CAGR Compounded Annual Growth Rate 36, 51

CAPEX Capital Expenditure 14, 16, 65, 73, 80, 81, 82, 83, 84, 91, 95, 96, 134, 168, 169, 171

CD Chromatic Dispersion 5, 11

CFP2-DCO 100G Form Factor Package 2 - Digital Coherent Optics 9, 71, 72, 73, 74, 76, 77, 78, 79, 80, 82, 83, 84, 168, 171

CLI Command Line Interface 11, 110

CNN Convolutional Neural Network 143

CRUD Create, Read, Update, Delete 38

CUT Channel Under Test 25, 98, 101, 102, 106, 167

DC Data Center 29, 31, 32, 33, 34, 35, 36

DOTN Disaggregated Optical Transport Network 109, 111, 112, 113, 169

DQN Deep Q-Network 89, 143

DSP Digital Signal Processing 4, 5, 66, 73, 101

DWDM Dense Wavelength Division Multiplexing 2, 7, 14, 17, 23, 66, 72, 109, 110, 115, 125, 137, 138, 139

EDFA Erbium Doped Fiber Amplifier 3, 5, 6, 9, 24, 25, 54, 59, 68, 80, 92, 138

EGN Enhanced Gaussian Noise Model 60, 138

ETSI European Telecommunications Standards Institute 14, 125, 128, 129, 130, 136

FEC Forward Error Correction 45, 57, 101, 102, 103, 107

FF-EGN Full Form Enhanced Gaussian Noise Model 46, 48, 49, 50, 51, 58

GN Gaussian Noise 21, 24, 26, 46, 47, 137, 138, 139

GRU Gated Recurrent Unit 99, 104, 106

GSNR Generalized Signal to Noise Ratio 24, 26, 29, 46, 47, 48, 49, 50, 57, 58, 59, 60, 61, 74, 76, 84, 89, 101

HTTPS Hypertext Transfer Protocol (Secure) 112, 128, 130

HTTPS/PQ Hypertext Transfer Protocol (Secure) Post-Quantum 127, 128

IETF Internet Engineering Task Force 37, 38, 127, 171

ILA In-line Amplifier 2, 7, 8, 10, 21, 24, 26, 28, 39, 73, 92, 117, 118, 119, 123, 134, 167

ILP Integer Linear Programming 20, 45, 52, 84, 85, 86, 87, 89, 93, 94, 95, 96, 134, 171

IP Internet Protocol 10, 11, 71, 72, 73, 83, 169

IPoWDM Internet Protocol over Wavelength Division Multiplexing 2, 9, 71, 72, 73, 78, 81, 83, 134, 136

ISRS-GN Inter-Sysmbol Stimulated Raman Scattering Gaussian Noise 74, 139

ITU-T International Telecommunication Union Telecommunications Standardization Sector 38, 73

IXP Internet Exchange Point 29, 31, 32, 33, 34, 35, 36

JSON Java Single Object Notation 37, 38

KME Key Management Entity 16, 18, 121, 126, 127, 129, 130, 131, 135

LASER Light Amplification by Stimulated Emission Radiation 3, 5

LSTM Long Short Term Memory 99, 104, 106, 108

MAE Mean Absolute Error 105, 106

ML Machine Learning 2, 17, 18, 21, 84, 99, 100, 102, 103, 104, 105, 108, 135, 136, 141, 169

MSA Multi-Source Agreement 39, 109

MST Minimum Spanning Tree 116, 118, 119, 123, 135

NIST National Institute of Standards and Technology 14, 127, 171

NLIN Nonlinear Interference Noise 8, 11, 23, 24, 25, 26, 46, 54, 56, 59, 60, 74, 89, 137, 138, 139, 167

OCSVM One Class Support Vector Machine 99, 103, 104, 105, 106, 107, 108, 143, 171

O-E-O Optical-Electrical-Optical 19, 71, 98

ODTN Open Disaggregated Transport Network 11

OLS Optical Line System 1, 2, 13, 17, 18, 24, 65, 71, 73, 79, 80, 81, 82, 83, 84, 97, 98, 100, 108, 109, 110, 171

ONF Open Networking Foundation 11, 38

OOPT Open Optical and Packet Transport 12

OPM Optical Performance Monitoring 98, 100, 101, 102, 103, 108, 110, 145

ORP Optical Receive Power 99, 101, 102, 103, 104, 105, 106, 107, 108, 143, 169

OSaaS Optical Spectrum-as-a-Service 18, 20, 97, 98, 100, 108, 143

OSNL Open SDN & NFV Lab, Berlin 111, 112

OSNR Optical Signal to Noise Ratio 23, 24, 26, 27, 60, 63, 101, 102, 167

OTN Optical Transport Networks 1, 2, 3, 4, 6, 7, 10, 11, 12, 13, 14, 17, 18, 19, 20, 21, 23, 24, 27, 28, 29, 31, 32, 35, 36, 37, 38, 39, 40, 43, 45, 46, 51, 54, 56, 58, 65, 67, 71, 73, 74, 76, 79, 84, 85, 95, 96, 97, 98, 99, 109, 110, 113, 115, 116, 117, 124, 125, 130, 133, 134, 135, 136, 167, 168, 171

PCE Path Computation Element 40

PKI Public Key Infrastructure 130

PMD Polarization Mode Dispersion 5, 11

PoP Points of Presence 2, 27, 28, 29, 30, 31, 57, 117, 118, 121, 134, 167

PQC Post-Quantum Cryptography 14, 20, 21, 115, 125, 126, 127, 128, 135, 136, 171

PSD Power Spectral Density 25, 137, 138

QAM Quadrature Amplitude Modulation 26, 45, 47, 66, 100

QKD Quantum Key Distribution 3, 14, 15, 16, 17, 18, 19, 20, 23, 41, 42, 115, 116, 117, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 135, 136, 167, 169

QKDN Quantum Key Distribution Network 14, 18, 20, 23, 41, 43, 115, 116, 117, 121, 124, 125, 129, 130, 131, 135

QoT Quality of Transmission 8, 19, 23, 24, 46, 57, 58, 68, 97, 100, 110, 112

QPSK Quadrature Phase Shift Keying 26, 45, 66, 100, 110

QSFP-DD Quad Small Form-factor Pluggable Direct Detection 72, 73

RCSA Routing, Configuration, and Spectrum Allocation 17, 19, 29, 30, 31, 40, 45, 46, 47, 52, 57, 65, 66, 67, 68, 69, 71, 75, 76, 83, 84, 134, 136, 147, 148, 167

REST Representational State Transfer 18, 39, 98, 125, 127, 128, 135, 136, 171

RL Reinforcement Learning 19, 84, 85, 89, 90, 91, 93, 94, 95, 96, 134, 141, 150, 171

RMSE Root Mean Square Error 105, 106

ROADM Reconfigurable Optical Add Drop Multiplexer 2, 8, 10, 21, 26, 28, 39, 92, 95, 98, 100, 109, 111, 112, 135, 167

RP Regenerator Point 56, 57, 58, 59, 60, 61, 62, 63, 168

RPC Remote Procedural Call 38, 39

RSA Routing and Spectrum Allocation 52, 58, 136

RSA-4096 Rivest-Shamir-Adleman 14, 115

SAE Secure Application Entity 118, 121, 122, 123, 124, 130, 131, 135, 169

SCI Self Channel Interference 25, 138

SDN Software Defined Networking 2, 3, 10, 11, 12, 16, 17, 18, 19, 21, 37, 38, 39, 41, 98, 100, 110, 111, 116, 125, 126, 129, 130, 131, 136

SKR Secure Key Rate 16, 17, 42, 43, 118, 119, 120, 121, 122, 123, 124, 167, 169

SLA Service Level Agreement 1, 65

SMI Structured Management Information 37

SNMP Simple Network Management Protocol 10, 37, 110

SNR Signal to Noise Ratio 8, 23, 24, 47, 49, 101, 102, 103, 137

SSFT Split Step Fourier Transform 25, 46

SSMF Standard Single Mode Fiber 3, 4, 5, 6, 15, 41, 42, 43, 65, 167

SVM Support Vector Machine 143, 144

TAPI Transport Application Programming Interface 11, 12, 13

TD Traffic Demand 56, 57, 59, 60, 61, 63, 168

TDFA Thulium Doped Fiber Amplifier 9, 68, 80, 92

TLS Transport Layer Security 98, 113, 128

TRN Trusted Node 16, 116, 117, 119, 121, 122, 123, 124, 130, 135

VPN Virtual Private Network 111, 113

WSS Wavelength Selective Switch 92, 95

XCI Cross Channel Interference 25, 60, 138

XML Extended Modelling Language 37, 38

XPDR Optical Transponder 109, 111, 112, 135

YANG Yet Another Next Generation 2, 10, 12, 13, 37, 38, 39, 110, 131

Chapter 1

Introduction

With the roll-out of 5G services for private mobile telephony users in several countries, communication network operators, as well as manufacturers, now focus their efforts to enable ultra-high speed and low latency services for other networking paradigms such as e-Health, Industry 4.0, connected mobility, etc. Meanwhile, as research and standardization of 6G has already begun, the network infrastructure to support low-latency and highly reliable applications needs to be planned in advance, Using pre-existing optical fiber infrastructure to interconnect these customer-facing front-haul networks enables operators to leverage the high throughput with limited investment offered by optical fibers and assure Service Level Agreements (SLAs) with their customers. As the demands in the access and metro region increase, assuming an increase in traffic throughput in long-haul country-wide backbone optical networks is valid. Over the past decade, many long-haul Optical Transport Networks (OTN) operators have upgraded their services from 10 Gbps direct detection to 40 Gbps and 100 Gbps coherent detection. However, it is seen in recent years these data rates can still lead to capacity exhaustion, forcing network operators to upgrade their networks [13].

Meanwhile, the European Commission guidelines state that network operators have to provide high bandwidth internet services to all residential and transportation users by 2025 [14]. Although plenty of dark fiber is available in western European countries, the continuous growth of demands would either force physical network infrastructure owners to upgrade their networks by adding additional dark fibers or invest in optical networking equipment vendors, who can offer both flex-grid as well as multi-band transmission devices [15]. Therefore, long-haul OTN operators need to seek ways to upgrade network capacity and find scalable and cost-effective solutions to monitor and control their networks.

Flex-grid OTNs allow for the deployment of flexible Bandwidth Variable Transceivers (BVTs) into the pre-existing Optical Line System (OLS) using the ITU-T recommendations [16]. Here, the spectral resolution of each grid is 6.25 GHz, allowing the optical digital

signal processors on board the BVTs to increase the symbol rate and reduce the spectrum used in the network. Network operators also have the option to undertake an Internet Protocol over Wavelength Division Multiplexing (IPoWDM) based terminal device upgrade, which can allow up to 400 Gbps colored optical interfaces directly mounted onto Layer-3 routers and switches. Such upgrades need to be undertaken only at the traffic Points of Presences (PoPs). However, the number of terminal devices does not scale well with the exponential traffic increase and further avenues to increase the capacity of the Dense Wavelength Division Multiplexing (DWDM) network need to be studied. As demands increase, OLS needs to be upgraded to further bands, i.e., L-Band and S-Band to maximize the throughput capacity on each fiber. With the technological maturity of multi-band optical transmission devices, specifically Reconfigurable Optical Add Drop Multiplexers (ROADMs) and In-line Amplifiers (ILAs), upgrading the network with such devices allows network operators to meet the traffic demands of the coming decades.

On the one hand, multi-band flex-grid OTNs provide flexibility and increase the network capacity, however, the transmission properties, as well as control plane structures, become more complex. Recently, Software Defined Networking (SDN) has emerged as an efficient way to manage heterogeneous devices in the network and allow quick provisioning of optical services in DWDM networks. Although many optical device vendors offer SDN-based control and data plane demarcation using their propreitary software, network operators and open-source consortium are pushing for the introduction of open Yet Another Next Generation (YANG) models into these devices. This step towards disaggregation of OTN devices creates a potential for network operators to develop open-source-based control plane solutions, thereby saving costs and avoiding vendor lock-in. Using open-source optical domain controllers, optical services can be easily added and managed, leading to the creation of pathways for integrating optical networking devices from multiple vendors into a given network operator domain.

Another important aspect to control in-service optical devices is to be able to monitor them continuously, and hence, use the performance monitoring data of BVTs to detect and possibly pre-empt a failure on a given device. Such early detection can be used to re-route existing mission-critical optical services onto different devices on a disjoint backup path, thereby avoiding loss of service. Since performance monitoring data can be extracted quickly from BVTs, Machine Learning (ML) based models can be leveraged to deduce useful information and detect anomalies in the network. The benefits of an ML-based approach are demonstrated and evaluated on BVTs under test, deployed in a long-haul live OTN network. This methodology showcases the applicability of ML-based modeling for real-time control of

OTNs. With this, the control aspects of this dissertation are evaluated before moving to the security related aspects.

Advances in semiconductor physics over the past decades have led to the development of quantum computers. Although still in an experimental phase, these computers can solve problems in linear time, which binary computers can solve in exponential time. One such problem is the factorization of large prime numbers, on which several software-based key encryption solutions are based. Moreover, using fiber-tapping methods, attackers could store this encrypted data and decode it after quantum computers are commercially available, leading to an immense security risk for owners of sensitive and mission-critical data.

To tackle these challenges, quantum-based cryptography techniques provide information-theoretic security. One such symmetric key generation method is called Quantum Key Distribution (QKD), which requires its own hardware components and associated dark fibers to exchange keys. The deployment of QKD devices requires a new network planning-based solution since the physical transmission characteristics of QKD differ from long-haul OTN devices. Finally, the keys generated by QKD devices need to be managed across the network and integrated into SDN-based optical domain controllers, thereby allowing for secure exchange of traffic across the network.

In this chapter, relevant system aspects of optical networks are discussed, including the physical transmission characteristics of flex-grid networks. Then, a brief overview of physical-layer network planning, disaggregated network control, and open-source initiatives for OTN management is provided. Further, QKD transmission and various terms related to quantum-safe optical networking are introduced. Finally, an outline of the dissertation, along with relevant publications is presented.

1.1 System Aspects of Optical Networks

Approaching from first principles, optical networking is based on photonic communication, which converts electrical signals into photons and transmits these photons over long distances using optical fibers.

Modern fiber optics transmission is dependent on several inventions over the past 60 years, beginning with the invention of a coherent optical source known as Light Amplification by Stimulated Emission Radiation (LASER). The first experiments to use an extremely thin transparent glass fiber as a waveguide for light was conducted in 1966 [17]. However, these transmission experiments could not extend the reach of the signal by more than a few kilometers. With the invention of Standard Single Mode Fiber (SSMF) in 1980s, long-haul transmission was made possible. Finally, the design of an Erbium Doped Fiber Amplifier

(EDFA) led to the possibility of amplification of optical signals, thereby leading to the replacement of copper-based cables by SSMF.

In telecommunication networks, optical communication is used to exchange terabits of information every second. To design and operate such a system, the underlying components, their physical limitations, and the overall laws of physics governing of optical transmission need to be understood first. Hence, the high-level design of an optical communication system is generally introduced, followed by the bands of operation as well as frequency grid used for optical channel transmission. Finally, an introduction into long-haul OTN and its associated components is also provided.

1.1.1 Optical Fiber and Communication System

Optical fibers are thin, light-weight transparent media, which can be used as a waveguide for a light source incident in it. It consists of two silica-based cylindric media, each having diameter in the order of micrometers and different refractive indices together. The inner medium is known as *core* and the outer medium is known as *cladding*. Since the core has a higher refractive index than the cladding, light traveling through the core undergoes total internal reflection, thereby propagating along the fiber with minimal loss of power and intensity. Broadly, optical fibers can be distinguished into two categories, that is, multi-mode and single mode. Multi-mode fibers allow light sources to propagate together with the same frequency at different phases, thereby creating a stationary wave. However, multi-mode transmission and fibers are susceptible to modal dispersion [17] and therefore cannot transmit data over large distances. The single mode fiber allows light to travel only on a single mode (i.e., fixed frequency and phase), thereby eliminating modal dispersion issues and allowing for long-haul propagation of the light source. The SSMF is the most commonly deployed and used single mode fiber and also is the only type of optical fiber used for network studies in this thesis. In recent times, different types of fibers like multi-core fibers, few-mode fibers, and hollow-core fibers have also emerged and claim to increase the per-fiber transmission capacity exponentially. However, due to ongoing research and standardization, they are yet to be deployed commercially.

The ability of SSMF to carry light over long distances with low loss is an inherent physical property that addresses the need of long-distance data communication. Since copper-based cables needed electrical repeaters quite often, they could not be used for long-haul communication at high data rates. This led to further development of coherent optical communication devices which could use modulators, lasers, and optical Digital Signal Processing (DSP) to encode binary electrical signals into optical signals to transmit and receive high-speed data across long distances.

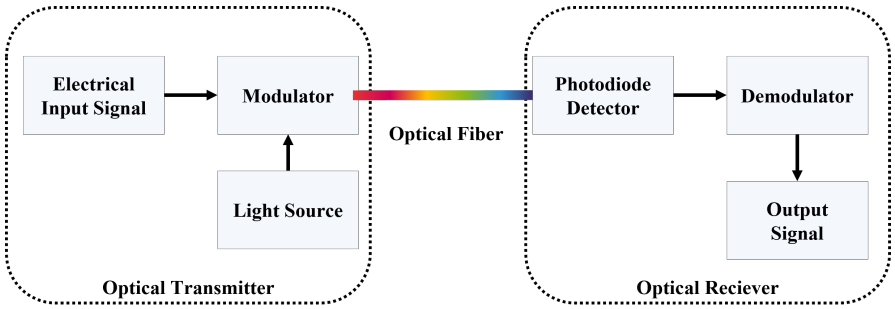


Figure 1.1: Block diagram of an Optical Communication System.

In Fig. 1.1, the electrical signal input to the optical transmitter is modulated onto a light source (LASER) at a given carrier frequency. Modulation formats are the methods used to encode binary information into light sources. These can either be amplitude modulation or phase shift keying. This modulated light signal is tuned to a frequency and propagated along the optical fiber. At the receiver side, a photo-diode converts the received optical signal into an electrical current, which is processed using a DSP to retrieve and reconstruct the output electrical signal. With recent advances in the development of optical DSPs, higher order amplitude modulation can be achieved on modulators at the transmitter side. Similarly, transmission effects such as Chromatic Dispersion (CD) and Polarization Mode Dispersion (PMD) can be compensated electronically at the receiver side.

1.1.2 Optical Communication Bands

Optical light sources (LASER) can be tuned to a provided wavelength for propagation through an optical fiber. For single mode fibers like SSMF, a cut-off wavelength is needed, beyond which modal dispersion occurs. The ITU-T G.652 standard, therefore set the lower limit of this cut-off limit to 1208 nm, from where the O-Band begins.

Since SSMF fiber has a high attenuation (water absorption) peak at 1383 nm, the E-band was created starting from 1360 nm upto 1460 nm. The 1460-1530 nm S-band is used for out-of-band monitoring and supervisory channels and the higher wavelengths of this band show attenuation characteristics similar to the conventional C-band.

From a data transmission perspective, the C-band, from 1530-1565 nm is of utmost importance, since it has been traditionally used to carry DWDM traffic. This is due to both the low attenuation regime as well as a relatively flat EDFA gain profile across all the wavelengths in the C-band. Here the exception in gain profiles is seen at 1530 and 1550 nm, which have a higher gain as compared to the neighboring wavelength. To counteract this phenomenon, gain flattening filters are used within the EDFAs.

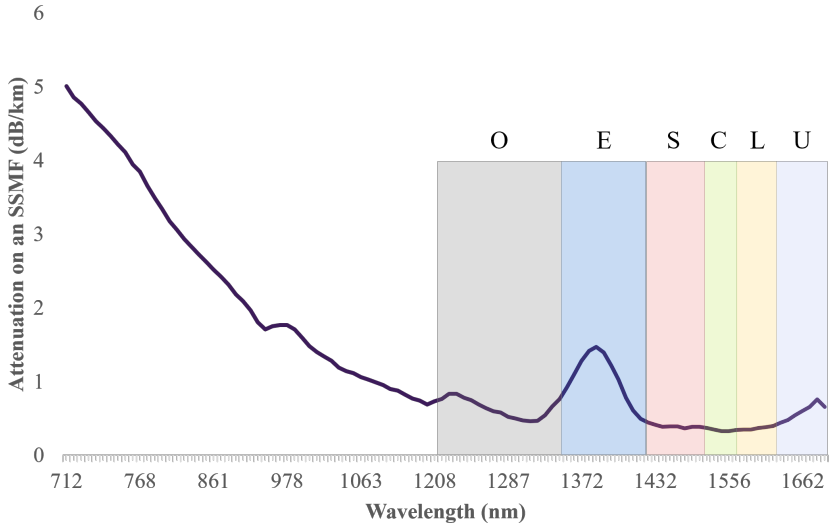


Figure 1.2: Optical transmission bands, of which S-band, L-band, and C-band are explored in this dissertation.

Similar to the C-band, the L-band (1565-1625 nm) shows minimal attenuation in SSMF, but needs a different design of EDFA to amplify the wavelengths in this band. Since the lower wavelengths of the L-band can be used to carry DWDM channels too, it is the first band which can be used in case network operators want to deploy multi-band operation in their networks. Fig. 1.2 summarizes the bands, along with the attenuation profile of a SSMF across the different bands.

In this dissertation, the various aspects of planning optical channels into the low attenuation regimes of SSMF, namely C-band, L-band, and S-band are studied and relevant strategies for upgrade into the different bands are discussed.

1.1.3 Dense Wavelength Division Multiplexing

The simplistic system of transmitting a single optical channel is not enough to meet the traffic demands between any given source and destination in commercial networks. Conceptually, optical communication systems can be applied to access, metro, and core networks (also known as OTNs). Although our analysis is restricted to long-haul OTNs, an introduction to different optical networks based on geographical and population needs is pertinent. Fig. 1.3 shows the demarcation between access, metro, and core networks. Access networks are defined as networks that connect several residential and business users to the nearest central office in their vicinity. These networks usually have a diameter in the order of a few kilometers. Metro networks aggregate the traffic of several smaller central offices and connect them to a core network location and usually have a transmission distance of 50-100 km. Finally, the

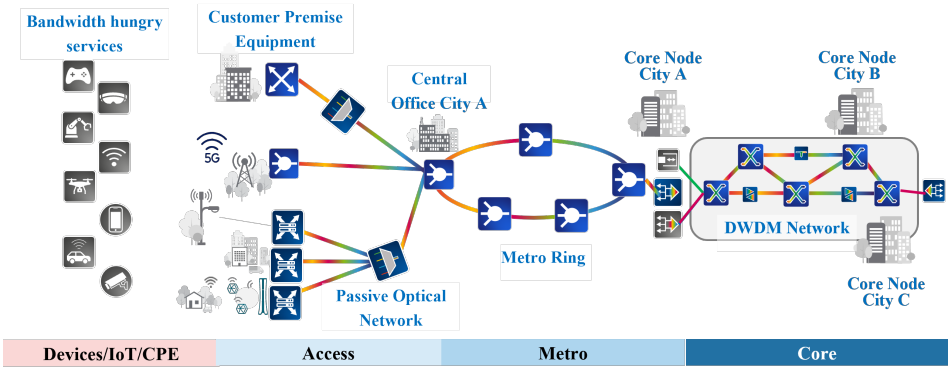


Figure 1.3: Optical network demarcations, based on geography and application. This thesis focuses on solutions for core network deployment.

aggregated traffic of each city or region at a core node is transmitted to other core nodes in different regions through the core network. The optical signal transmission distance can range from a few hundred to a few thousands of kilometers. The dimensioning of each of the networks and the traffic exchanged between different kinds of networks are interdependent on each other. Since access and metro networks need not cover large distances, their requirements can be satisfied by passive optical networks which do not contain any active ILAs between the traffic add/drop nodes. However, long-haul core networks need to optically amplify signals over large distances, and therefore use ILAs in case the distance between two core locations is greater than 80 km.

To allow transmission of several optical channels over long distances using optical fibers, the concept of OTN transmission was introduced. OTNs were traditionally deployed over the long-haul network, with each optical channel between traffic generating nodes capable of 100 Gbps data rate. A frequency domain representation of two light sources between Nodes A and B is shown in Fig. 1.4.

1.1.4 Optical Networking Components

The OTN, in principle, consists of a collection of DWDM systems. A DWDM system is defined as an optical communication system that can multiplex several high data-rate optical channels onto a single optical fiber. The three major components in an OTN are illustrated in Fig. 1.4 and can be described as follows:

1. BVT - These devices aggregate client traffic from core routers (10-100 Gbps) onto a single **lightpath** connecting the source and destination. Each lightpath is tuned to a wavelength, as defined by the DWDM spectral grid in the ITU-T G.694.1 standard. For all the fibers which carry a given lightpath between the source and destination, a central

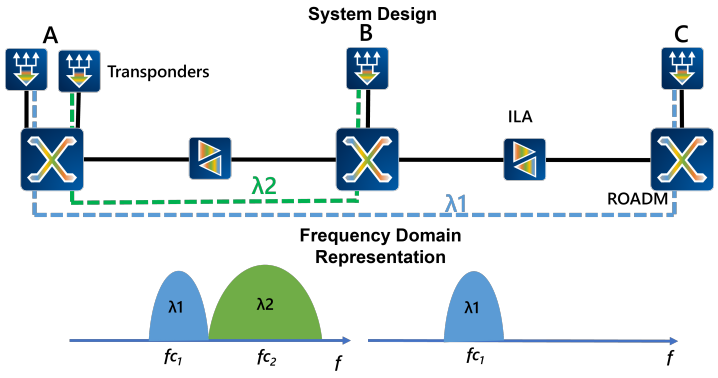


Figure 1.4: Optical Line System between three nodes, with its components along with frequency domain representation of two lightpaths λ_1 and λ_2, with central channel frequencies of fc_1 and fc_2 respectively.

wavelength is reserved for the given lightpath. Each lightpath can offer a data-rate anywhere between 100-600 Gbps, based on the BVT's minimum receiver sensitivity Signal to Noise Ratio (SNR). In recent times, advances in digital signal processing have led to the Several transponders operating simultaneously can lead to multiple optical channels deployed on different wavelengths along the same fiber. These channels cause an interference among themselves due to the four-wave mixing phenomenon, leading to the generation of Nonlinear Interference Noise (NLIN). NLIN is a major noise component that leads to Quality of Transmission (QoT) degradation, thereby limiting the traffic carrying capacity of each BVT.

2. ROADM - Lightpaths in the network need to be routed between source and destinations. For this purpose a ROADM is used. The latest technologies in ROADMs allow a colorless, directionless, and contention-less mode of operation. Using a ROADM, several lightpaths can be added, passed-through, or dropped at the source, intermediate, or destination nodes in the network respectively.

3. ILA - Distances between cities in a country are often in the order of several hundreds of kilometers. A lightpath, consisting of photons, can travel up to 80 km, before transmission related losses start occurring. Therefore, at multiple locations between two ROADM nodes, an ILA node is required. An ILA can amplify the optical signal, thereby allowing it to travel longer distances without the need for opto-electronic regeneration and reshaping or re-framing of the signal. However, each amplification adds Amplified Spontaneous Emission (ASE) noise onto each lightpath, further degrading the QoT of the lightpath.

1.1.5 Capacity Upgrade Solutions

As standardization bodies and researchers have already begun efforts to define latency and bandwidth requirements for services in the 6G ecosystem, core network operators need to plan a budget for network investments to have sufficient capacity provisioned when the need arises. There are several commercial offerings available from optical transport network vendors, which can be utilized by network operators to scale their capacities, including optical terminal equipment upgrade [18].

Another way of increasing capacity is to use IPoWDM pluggables which can convert 100, 200, and 400 Gbps ethernet client rates directly to 100, 200, or 400 Gbps optical signals. Specifically, for transparent long-haul applications, the OpenROADM-based 100G Form Factor Package 2 - Digital Coherent Optics (CFP2-DCO) modules provide an alternative to BVT-based deployment [19]. These modules can connect directly to a high-speed switching router which can further be connected to the Open Line System (OLS).

At the OLS layer, transparent IPoWDM has emerged as a cost-effective competitor to the BVT-based deployments; however, higher costs related to coherent optics-enabled routers, and the versatility of BVTs must also be considered in an upgrade decision. Moreover, as shown in Table 1.1, multi-haul CFP2-DCO modules can achieve a maximum of 16 QAM and 64 GBaud configurations. An 'a-la carté' deployment of IPoWDM and BVTs might result in the best of both worlds and eventually lower capital expenditure than pure deployments; although network operations, maintenance, and logistical cost of such an approach still need to be adequately studied.

Sl. No.	Technology	Data Rate (Gbps)	Modulation Formats	Baud Rate (GBaud)
1	BVT	100-600	QPSK-64QAM	28-72
2	CFP2-DCO	100-400	QPSK-16QAM	28-64

Table 1.1: Comparison of differences between BVTs and CFP2-DCOs in terms of optical performance.

Network operators need to evaluate these terminal upgrade solutions to make strategic investment decisions for a future-proof provisioned network capacity. Moreover, as wide-band low noise figure EDFA and Thulium Doped Fiber Amplifier (TDFA) technologies mature, these may soon be included in commercial network deployment. The question therefore arises, as to how soon must a network operator deploy a multi-band solution. Adding additional bands is a cost-intensive exercise, due to the addition of band splitters and combiners at each amplification point in the network. However, several studies have shown that such an investment is bound to be cheaper than lighting additional dark fibers while offering similar growth in carried traffic [20, 21].

1.2 Software Defined Networking for Optical Networks

As a concept, SDN has been researched upon since the beginning of the millennium. Considering its success in simplifying network control functions in the Internet Protocol (IP)-layer devices like packet switches and routers, this concept has recently been applied to optical networks, specifically OTNs as well.

SDN invokes the separation of the control and data plane within a networking device. This allows for the creation of different abstraction layers for different interfaces of the device, thereby leading to a distributed software architecture [22]. The SDN architecture consists of three layers, namely, an orchestration layer, a control layer, and a device layer. Layers interact with each other using Application Programming Interfaces (APIs), which are called either as northbound interfaces if they connect the control layer to the orchestration layer or as southbound interfaces if they connect the control layer to the device layer.

This section provides an overview of the SDN architecture by first defining the terms used as well as the functionality of the different parts. Then it delves into SDN solutions specifically for optical transport networks, and finally discusses the concept of open-source disaggregation for optical transport networks.

1.2.1 SDN terminologies and functions

A simplified SDN architecture applied to optical networks is depicted in Fig. 1.5. At the lower-most layer, i.e., device layer each OTN device like BVT, ROADM or ILA can be defined by a set of data models which characterize the behaviour and functionality of the device. These data models are known as YANG models and are used to define various device parameters. The device parameters can be configured using Southbound APIs from an SDN domain controller, which has a global view of all the devices connected to its domain. The connected devices can be accessed by different protocols like NETCONF, RESTCONF, or Simple Network Management Protocol (SNMP), which allows for programmatic provisioning and control of the devices.

At its core, the SDN controller is a network management system which can hold information about every device in its domain, as well as provide fault, configuration, accounting, performance and security functionality in the network. Some SDN controllers have additional features like path calculation and load balancing. The SDN controller also exposes Northbound APIs which can be used for business use-case based applications like multi-domain network orchestrators, network monitoring tools, online network planners, etc. The Northbound APIs are usually handled by the RESTCONF protocol, which allow applications to request changes into the devices via the controller or request information about the devices

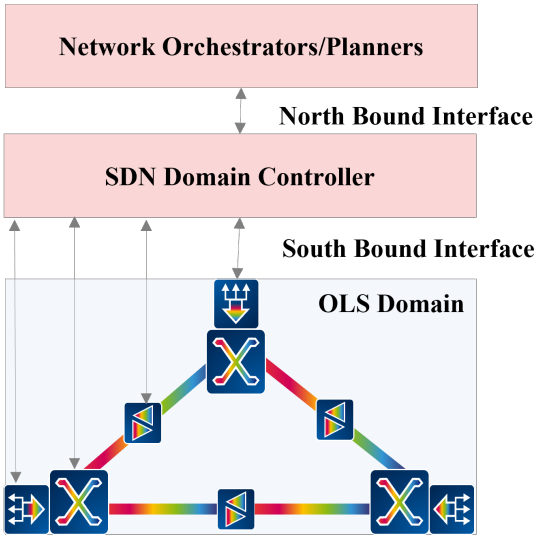


Figure 1.5: Three node OTN connected to an optical domain controller, which is further connected to a network orchestrator.

connected to the controller. Such an approach simplifies the previous methods of provisioning and monitoring each networking device manually using the device's Command Line Interface (CLI) and reduces chances of human errors which may lead to service degradation.

1.2.2 Open-source SDN Controllers for optical transport networks

For IP-based SDN south-bound interface, OpenFlow is a widely used open-source protocol. However, optical networking devices have many physical parameters which are unique to fiber optics transmission and OpenFlow which was primarily designed for packet-based networks needs to be extended significantly to meet the configuration and monitoring needs of optical transport network operators. The challenge faced by the optical networking SDN community is to create a data model which is abstract enough to be controlled by higher layers while keeping a track of optical communication parameters like NLIN, PMD, CD, etc.

To cater for the needs of OTN operators, two open-source projects have been used extensively both in research as well as commercially. The first is Open Disaggregated Transport Network (ODTN) which is based on Open Networking Foundation (ONF)'s ONOS SDN controller [23]. The second project is TransportPCE, which is built on Linux Foundation's OpenDayLight SDN controller [24]. Both ODTN and TransportPCE are operator-led consortiums and seek to drive down network costs by using open data models and APIs across different optical networking devices. The main difference between the two controller projects is that while ODTN has been developed together with Transport Application Programming Interface (TAPI) data model and APIs, TransportPCE has primarily adopted OpenROADM-

based models across its whole platform. Both the projects are under constant development and have not yet been deployed onto live networks. Recently, a third consortium called Open Optical and Packet Transport (OOPT) under the Telecom Infra Project has emerged. OOPT till date has preferred to use ONOS-based SDN controller with TAPI data models for its first experimental demonstration. However, OOPT's network planning tool called GnPY is already integrated into TransportPCE as a physical layer path verification tool [25]. This is one of the first steps being taken to combine the two solutions and aim for completely interoperable data models in a disaggregated optical transport network.

1.2.3 Towards Disaggregated Optical Transport Networks

Disaggregation, by definition, means to separate an object into its constituent parts. In OTNs, disaggregation points to the efforts of creating standardized inter-operable data models which would allow network operators to deploy large-scale networks without having to either buy all the devices from the same vendor, or depend on vendor-specific network control software. According to Giorgetti *et. al.*, disaggregation can be broadly divided into two parts, namely, vertical and horizontal disaggregation [26]. Vertical disaggregation aims at separating the control plane from the data plane, as explained in the previous section. Horizontal disaggregation aims to create models for each optical networking device, thereby allowing network operators to select the best devices in a vendor-agnostic fashion.

At first glance, complete disaggregation will help drive down costs for network operators. However, the technological maturity of data models is currently low and different consortia come up with their own requirements, leading to an absence of standardization. In such cases, the integration costs of a disaggregated optical transport networking device into a fully disaggregated system will be high. From a data model perspective, OpenConfig and OpenROADM define YANG models for optical transport networks. OpenConfig is highly abstract and invites modifications by vendors to make its YANG models more robust. However, it is used extensively for connecting optical terminal devices like BVTs. OpenROADM has its own model at each vertical disaggregation model, namely, device, network, and service model. Each has its own set of YANG models which can be manipulated to create, monitor, and delete optical services. OTN device vendors now have a choice to support YANG models from either one or both. However, full compliance to either OpenROADM or OpenConfig would require restructuring of the entire software stack while ensuring back compatibility to legacy software.

This development intensive effort is currently not economically beneficial, since both OpenROADM and OpenConfig consortia are open to changes and keep releasing newer versions. However, since more network operators request for OpenConfig or OpenROADM

support, device vendors support specific YANG models based on the customer needs. This approach has led to the development of the concept of a partially disaggregated optical transport network system. In a partially disaggregated optical transport network, devices from

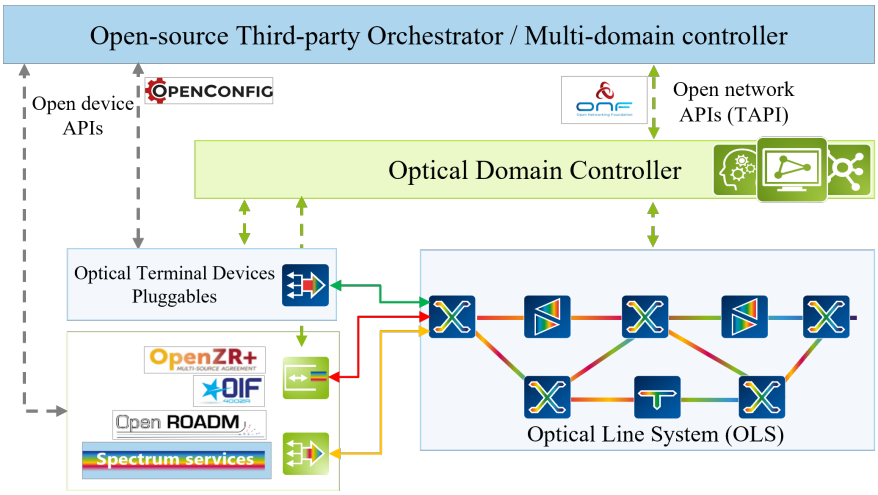


Figure 1.6: Example architecture of a partially disaggregated OTN.

different vendors offer basic support for topology discovery, provisioning, and performance management via open-source data models and southbound APIs. All other functionalities like alarm management, fault localization, restoration, etc., are handled by the vendor specific YANG models, connected directly to the vendor's domain controller. This offers network operators greater control over the networks via the vendor-specific software, as well as flexibility to implement cross-domain network applications.

Fig. 1.6 shows an example of how optical terminals as well as OLS can be connected to an open-source orchestrator. Since optical terminals and pluggables are easier to manage via open interfaces, they can either connect to the vendor's proprietary optical domain controller using proprietary NETCONF/RESTCONF and YANG models or using open-source YANG models such as OpenConfig directly to the open-source orchestrator. The OLS in a partially disaggregated network communicates only with the proprietary optical domain controller which in turn exposes TAPI based north-bound interfaces to communicate with open-source orchestrators. In this way, vendors can offer network operators a way to use open-source abstract models to perform highly specific operations on the OLS. From the current software landscape of vendor-specific optical domain controllers, mediators for each of the different open-source APIs and YANG models (TAPI , OpenConfig, OpenROADM) need to be implemented. Although a white-box OLS deployment with custom network control software is envisioned theoretically, from a practical perspective partial disaggregation

provides network operators the software flexibility to deploy solutions from multiple vendors thereby minimizing Capital Expenditure (CAPEX).

1.3 System Aspects of Quantum Key Distribution

Quantum threat is well recognized in optical networks since strong encryption over fiber is a key requirement of many applications e.g., government operations and critical-infrastructure operations like electricity or gas. While data encryption using symmetric-key cryptography (e.g., AES-256) can survive with enlarged key size, key exchange or authentication schemes using public-key cryptography (e.g., Rivest-Shamir-Adleman (RSA-4096)) are in danger. To defeat this threat, two approaches are usually taken in the industry. One is to establish a Quantum Key Distribution Network (QKDN) and supply encryption keys to data encryptors via a secure interface. The other is to implement Post-Quantum Cryptography (PQC) software suites directly on data encryptors and run a PQC key exchange protocol as well as data encryption on the same machine. Although both approaches have been investigated intensively, industry is unpleasant to deploy these methods widely on existing OTN mainly due to the lack of standardized technology. National Institute of Standards and Technology (NIST) is driving the standardization of new PQC algorithms that could withstand quantum attacks. The European Telecommunications Standards Institute (ETSI) and the International Telecommunication Union (ITU) have put efforts into the standardization of QKD and its application on telecommunication networks at scale [27, 28]. Hence, the industry prefers to open the possibility to take both approaches until the standardization processes are finalized.

However, deploying quantum-secure solutions in OTN is not a simple task in practice. QKD suffers from a limit of distance and the requirement of expensive hardware equipment. PQC solutions are flexible and easy to apply, nevertheless, they often require a powerful processor and relatively large memory resources that most of the embedded platforms in OTN may not have. In case NIST decides to standardize more than one PQC finalists, platforms need to have the capability to accommodate multiple PQC cipher suites, which may cause a significant change in the hardware architecture e.g., processing power and memory. A commercial PQC chipset or custom hardware accelerator might be another option but the time to market for such products are long as this approach is challenging to realize on an industrial scale.

From a network planning, deployment, and control perspective, QKD is more interesting, since feasibility studies need to be conducted and routing optimizations can reduce the cost of deploying such a network in parallel with the existing DWDM infrastructure. This thesis, therefore, focuses purely on QKDN. This section first provides a bird's eye summary of

achieving quantum cryptography using a QKD system. Then, it briefly discusses the details of a key exchange protocol which is used in commercial QKD devices.

1.3.1 Quantum-Safe Transmission using QKD

As discussed earlier, computational security offered by software-based cryptographic techniques like AES-256 on encrypted data transmission does not guarantee information-theoretic security. A communication channel carrying encrypted data is defined as information-theoretic secure if the attacker (Eve) cannot reconstruct the original message from the encrypted message, despite having infinite computational power. Another method of achieving quantum-safe communication between two nodes is using a QKD system.

In a QKD system, the key is exchanged over a dark SSMF which is used to transmit single photon qubits. A qubit, also known as a "Quantum Bit", can encode information based not only on the binary bits (0 or 1), but also using a photon's polarization state. Each qubit is encoded onto a photon using a combination of bits and polarization state. This combination is known as basis, and each qubit is launched into a dark fiber with no other channels present in it. Dark fibers are needed since the qubits possess low energy and are easily susceptible to information loss, especially as the distance increases. At the receiver side, a highly sensitive photodiode detects the series of qubits and extracts a raw key out of it.

To generate a secure key between Alice and Bob, the following four steps are needed, namely:

1. Quantum Transmission - Using a QKD protocol, a series of random qubits are transmitted over a dark fiber. In such a protocol, Alice first creates a string of qubits using a randomly selected basis. Bob also creates a random basis and uses it to measure the received qubits. Meanwhile, the bases are broadcast over the public channel by each party, such that Alice receives Bob's basis and vice-versa.

2. Sifting - The bases are compared and the non-matching bases are removed, leading to a unique measurement basis at Alice and Bob, enabling them to obtain the same key when they use the same measurement basis. The process of discarding the bits which were generated from different bases is known as sifting. The bits remaining after basis creation is known as the sifted key.

3. Error correction - In this step, Alice releases some additional information to Bob over the public channel, which can be used by Bob to correct erroneous qubits and also calculate the qubit error rate. Since the public channel is used, the information exchange for this step should be as minimal as possible.

4. Privacy Amplification - Since some information may have leaked to the eavesdropper (Eve) during the previous steps, the sifted and corrected key is compressed. This compression is done by a security proof, which first bounds the amount of information leaked and then relates it to the amount of compression needed to obtain a secure key. Each QKD protocol has its own technique to perform privacy amplification, which eventually results in a secure symmetric key.

The secure key generated at Alice and Bob is symmetric and is stored in a key management server, which can be managed using SDN techniques. Both the QKD device, as well as the key management server need to be located in highly secure location which denies physical or network access to all, barring the IT security personnel operating the devices. These locations, within the central office or core node locations are known as Trusted Nodes (TRNs).

Fig. 1.7 shows the QKD architecture example which is used in this dissertation. Each QKD device requires one dark fiber pair (quantum channel) and one classical fiber pair (service channel) [27]. Since QKD is a symmetric key distribution system, the keys need to be stored symmetrically in the Key Management Entitys (KMEs) at each TRN. Therefore, a Secure Key Rate (SKR) is used as a performance metric to measure the rate of information exchange between a pair of QKD devices. SKR is highly dependent on not only the fiber spans parameters like length and attenuation but also on the QKD device itself, where each device has its own parameters, i.e., QKD protocol, pulse repetition rate, dark count, etc. [29].

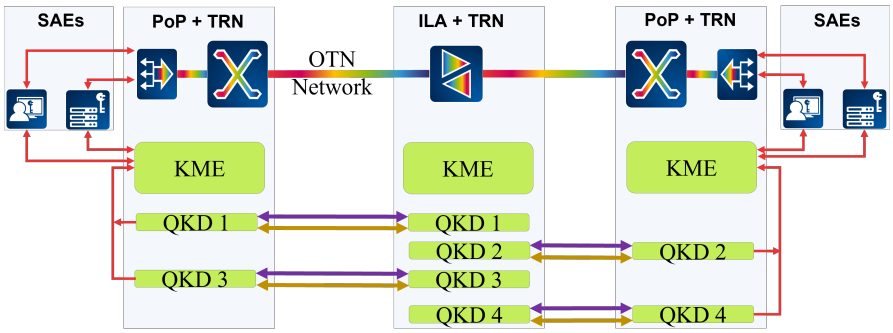


Figure 1.7: Example of quantum-safe traffic between two trusted nodes in the network using QKD.

Since the SKR is highly dependent on the QKD device, the protocols used, as well as the length of the fiber span between every Alice-Bob pair, a trade-off between number of TRNs in the networks versus the key-generation rate of the overall network is envisioned. Specifically, if the required key generation rate in the network is higher, more trusted nodes need to be deployed and at shorter distances. This is not only a CAPEX intensive exercise, but also adds to the operating expenditure, since TRNs need additional physical security and access measures. The increase in operating expenditure is also due to the additional security

measures required to convert a node in the network into a trusted nodes. In this dissertation, the relation of network deployment decisions to the QKD SKR is studied in detail.

1.4 Research Questions, Contributions and Outline of Dissertation

This dissertation addresses the planning and migration of DWDM optical transport networks towards flex-grid, multi-band and quantum-safe transmission, while investigating open-source control and monitoring mechanisms.

For achieving this, heuristics and meta-heuristics to solve the Routing, Configuration, and Spectrum Allocation (RCSA) problem for planning the physical layer in flex-grid OTNs are proposed. Then, network planning simulations are conducted to evaluate the effect of using different optical terminal architectures, as well as the effect of multiple bands in the network. Once the physical layer devices for future DWDM transmission is placed, a methodology and proof-of-concept to integrate open-source device abstraction models into an open-source optical domain controller, thereby successfully demonstrating the usage of such a solution to add, delete, and monitor lightpaths in the network. Further, using data from a production DWDM network, an SDN-based fault detection solution is created. This solution relies on ML models to detect failures in the network. Finally, methodologies to deploy QKD networks in existing DWDM networks, along with key storage and management techniques are investigated.

Therefore, the research question itself can be broken down into the following parts, each of which is addressed in the authors publications as indicated:

1. **RQ1** - Can commercially available flex-grid BVTs be deployed in optical transport networks to improve the network's spectral efficiency? If yes, can we optimize such deployments over multiple periods, such that the network throughput is not overprovisioned or underprovisioned?

   C1 - To answer RQ1, firstly, a heuristic for BVT placement in optical transport network is developed [3]. Then, using a combinatorial optimization-based solution, the placement of BVTs in a multi-period physical layer network planning scenario is optimized [4]. Further, a solution for translucent optical transport networks with a regenerator placement heuristic is proposed [5]. All these solutions address the general RCSA problem in flex-grid optical transport networks.

2. **RQ2** - Can C-band OLS in optical transport networks be upgraded to multi-band systems? If yes, is there a way to plan the migration to multi-band systems?

C2 - Since more-than-C-band OLS are already being commercially offered by OTN vendors, as a first step, a C+L-band study is undertaken to explore the benefits of a multi-band optical transport network deployment [6]. The relation between different optical terminals and upgrades to L-band and S-band is then analysed in detail [1]. Finally, a reinforcement learning based solution is provided to find the optimal number of bands and fibers, which will be added into the network over the given planning period [2]. RQ2 is answered with these works.

3. **RQ3** - Can optical transport network users detect failures in the network and devices using performance monitoring data? Can such devices, along with multi-band OLS systems be used to create, monitor, and delete optical services using open-source network control solutions?

   C3 - To answer RQ3, the Optical Spectrum-as-a-Service (OSaaS) use-case is first explored. In OSaaS, optical terminal devices like BVTs are operated by the end-users, giving them more flexibility over their optical services. However, OSaaS users have no OLS information . With the help of ML, a soft-failure detection methodology for OSaaS users is proposed and tested on BVTs deployed on a live production network [7]. Further, a proof-of-concept of deployment of an open-source SDN controller-based solution along with an online network planner and orchestrator is demonstrated, which is able to plan, add, monitor, and remove optical services from the connected devices [8].

4. **RQ4** - Can QKDN be deployed in conjunction with the current optical transport network infrastructure? If yes, can an SDN control of QKDN along with other quantum-safe encryption devices be envisioned?

   C4 - RQ4 deals with quantum-safe solutions' network planning, deployment, and control aspects. Here we first conduct a comparison of different techniques to reduce the number of trusted nodes while deploying long-haul QKDN across different network topologies [9]. Then, we look into a key delivery methodology by first proposing a cryptographic module capable of interacting with a KME using standardized Representational State Transfer (REST) APIs [10]. Finally, we provide a scalable SDN-controlled quantum-safe key management solution which can be deployed on OTNs with minimal implementation overhead [11].

Fig. 1.8 shows an overall outline of the thesis chapters, which are split according to the research questions and their specific contributions.

Chapter 1 deals with the introduction of the dissertation and consists of the system aspects of optical and QKD networks, along with the research questions and the thesis

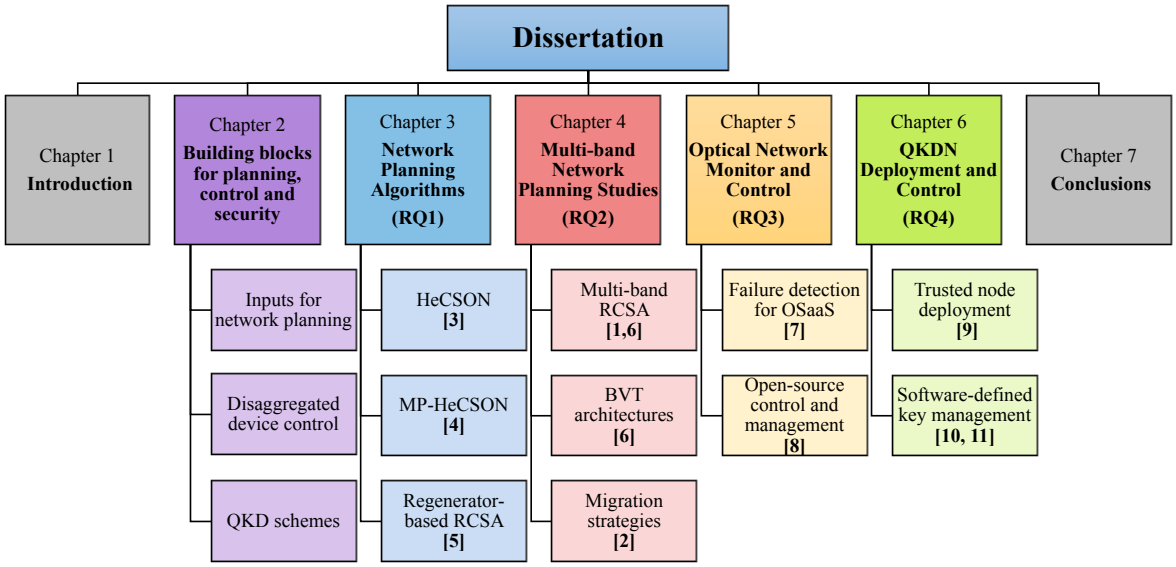


Figure 1.8: Visualization of the dissertation and the research questions.

outline. Chapter 2 begins by introducing concepts specific to the QoT in optical networks in Section 2.1. Then, network planning concepts are dealt with in Section 2.2, which will be used extensively in subsequent chapters. In Section 2.3 we dive deeper into the traffic models which was developed alongside the planning algorithms. Further, an open-source SDN controller used in this work, as well as a comparison of different QKD algorithms are discussed in Sections 2.4 and 2.5 respectively.

Chapter 3 addresses **RQ1** using three major contributions, namely, a heursitic for BVT configuration selection, named as *HeCSON* [3], a meta-heuristic for planning BVT deployment in a multi-period long-haul OTN deployment [4], and a regenerator-based RCSA heuristic developed to compare to the effects of deploying Optical-Electrical-Optical (O-E-O) regeneration in BVT-based networks [5].

Chapter 4 answers **RQ2** by extending the previous chapter's work into a multi-band scenario. In Section 4.1, we first undertake a network study to compare the regenerator-based options for C+L-band deployments for different long-haul OTNs [6]. Then, we compare different BVT architectures in Section 4.2, and their impact on multi-band activation for three differently sized long-haul OTNs [1]. Finally, in Section 4.3, we discuss a migration strategy for different networks, using an optimal ILP and a Reinforcement Learning (RL)-based solution [2].

Chapter 5 addresses **RQ3** by proposing a network monitoring and failure detection method for OSaaS users [7]. Then, a proof-of-concept for open-source disaggregated network management is presented, using which, optical services can be monitored and managed [8]. Chapter 6 deals with the QKD aspects of this dissertation and answers **RQ4**. Section 6.1 deals with the network planning aspects of QKDN deployment [9]. Section 6.2 deals with the control and management aspects of QKD and PQC quantum-safe keys and presents a proof-of-concept and a methodology to deploy such a solution on existing OTNs [10, 11]. In the end, Chapter 7 concludes this work and provides a holistic outlook on the challenges and opportunities for future work.

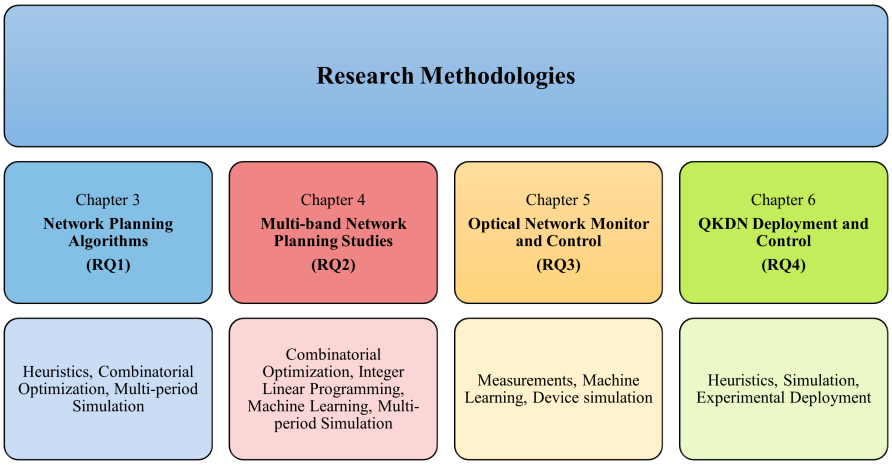


Figure 1.9: Research methodologies used in this dissertation.

To answer the four research questions, several research methodologies have been used in each of the contributing chapters of this dissertation. These research methodolgies are shown in Fig. 1.9.

In Chapters 3, 4, and 6, we use heuristics and combinatorial optimization to build network planning algorithms. Combinatorial optimization are a sub-field of optimization, which use a combination of heuristics and Integer Linear Programmings (ILPs) to arrive at a solution without resorting to brute-force or an exhaustive search. Most of the network planning problems in this dissertation are solved using this approach.

To verify network planning algorithms, a simulation environment which can transcend into multiple planning periods is needed. Therefore network studies use simulations to mimic network-wide effects of the planning algorithms. In this dissertation, Chapters 3, 4, and 6 simulate long-haul OTNs to derive network-agnostic solutions to the network planning and deployment problem.

In Chapters 4 and 5, different machine learning algorithms are used to solve tasks which are either complex to be solved by combinatorial heuristics, or need extremely quick results. This

dissertation compares machine learning-based algorithms with mathematical counterparts as a baseline, in order to highlight the actual benefits of using such algorithms to solve network planning and monitoring problems.

Chapter 5 also introduces a measurement setup for long-term data collection from BVTs deployed on a live long-haul OTN. This methodology ensures efficient transfer of monitoring data between multiple stakeholders and is used as a baseline for future streaming telemetry topics. Similarly, device simulations are also implemented to mimic real-lief BVTs, ROADMs, and ILAs, leading to a device level simulation of the Nobel-Germany [30] network using Docker containers. Finally, for SDN controller and PQC-based demonstrations in Chapter 6, experimental deployment of proof-of-concept code on optical network vendor devices is carried out. This shows the versatility of such devices to be modified quickly in order to cater for the need of quantum-safe optical transmission in the near future.

The author's other works, though sporadically referenced in this dissertation, do not form part of the contributions of this work. Furthermore, the appendix covers essential mathematical concepts, including Gaussian Noise (GN)-models for comprehending physical layer optical transmission, as well as the theoretical aspects of ML models for OTN fault detection and migration strategies. The appendix also provides pseudo-code for the algorithms used in RCSA and multi-band migration strategy.

Chapter 2

Building Blocks for Network Planning, Control, and Security

In Chapter 1, we provided a brief introduction to the system aspects of OTNs and QKDNs. Now, we delve into knowledge that is needed as a pre-requisite to understand the intricacies of OTN and QKDN planning and control. In the subsequent sections, we improve our understanding of noise that occurs during coherent DWDM transmission, and how NLIN plays an important part in selecting a configuration for BVTs in a long-haul OTN.

Then, we proceed to explain the terminologies as well as methodologies to generate topological inputs for physical layer-based network planning studies. One of these inputs is the traffic demands, which can be addressed using traffic models. We present the evolution of the traffic model used in our work across different contributions. We further discuss data models, protocols, standards, and open-source initiatives for OTN control. Since the implementation of an end-to-end service management proof-of-concept is a contribution of this dissertation, we discuss the software architecture of OpenDayLight TransportPCE combined with OpenROADM data models. Characteristics and transmission reach of a QKDN are dependent on the type of QKD device, as well as the type of QKD algorithm being used. Here, we compare different algorithms and settings for a QKDN and discuss the advantages and disadvantages of various approaches.

2.1 Physical layer Quality of Transmission

2.1.1 Optical and Generalized Signal to Noise Ratio

SNR usually refers to the ratio of the power of a transmitted signal to the power of the noise accumulated during transmission. A higher SNR refers to a better QoT. In optical transmission, especially in experiments, the Optical Signal to Noise Ratio (OSNR) is measured, where a

probing oscilloscope measures the ratio of the optical channel power to the noise at a reference bandwidth of 12.5 GHz[1]. However, since this dissertation deals with wideband network simulations, we focus on the calculated OSNR value[2], which is also termed as Generalized Signal to Noise Ratio (GSNR) [31]. For every optical channel propagating in an OLS, we can calculate the QoT using the GSNR which is defined as follows:

$$GSNR = \frac{P_{ch}}{P_{ASE} + P_{NLIN}} \tag{2.1}$$

where, P_{ch} is the power of the optical channel as it enters into the fiber, P_{ASE} is the ASE noise power. Since ASE noise power scales linearly with ILA, the closed-form equations provide an accurate estimate [32]. P_{NLI} is the NLIN power and is approximated by using closed-form GN models [33].

The GSNR of an optical channel provides information about its capability for error-free transmission. However, lightpaths with different channel bandwidths, as shown in Fig. 1.4 can cause inter-symbol and inter-channel interference, leading to QoT degradation. Therefore, network planning algorithms need to cater for both linear noise and NLIN in their calculations, before they can guarantee that a candidate lightpath would provide error-free transmission upon deployment. A brief description of both follows.

2.1.1.1 Linear noise in OTNs

Long-haul OTN transmission is dependent on ILAs. These devices optically amplify the lightpaths which are input to it. For long-haul transmission, a variable-gain EDFA is used, since its gain characteristics can be modified to amplify both C and L-band optical channels. For an EDFA to amplify the power of an input optical signal, its gain at a particular frequency/wavelength needs to be equal to the loss experienced by the optical signal at that wavelength while traversing through the previous span. In wide-band EDFAs, the power over multiple optical channels also needs to be equalized in order to avoid a gain tilt. However, with every EDFA traversed, this gain in power also adds a noise known as the ASE noise. In an amplifier chain, the ASE noise added by each amplifier can be given as described by Zyskind *et. al.* [34]:

$$P_{ASE} = 2h\nu \cdot \delta\nu \cdot n_{sp}(G - 1) \tag{2.2}$$

where h is the Planck's constant, ν is the central channel frequency of the optical channel under test, $\delta\nu$ is the optical bandwidth (0.1 nm conventionally), n_{sp} is the spontaneous emission

[1]Optical bandwidth and channel spacings are traditionally measured in wavelengths, i.e., 12.5 GHz is equivalent to 0.1 nm. In this dissertation, we use the frequency (GHz) notations for channel frequencies and bandwidths.

[2]The only exception to this rule is seen in Section 5.1.4, where OSNR and SNR is read out from an optical terminal device.

factor, determined by the inversion of Erbium ions in the EDFA, and G is the gain of the amplifier. Since the ASE noise is treated as a linear system, the noise after N similar amplifiers in an amplifier chain is given as $N \cdot P_{ASE}$. In case the gain and noise figure characteristics of each amplifier are different, the P_{ASE} value of each EDFA can be summed up to find the linear noise of the whole link.

2.1.1.2 Non-linear interference noise in OTNs

NLIN has been under research since the early 1990s with several existing analytical solutions to mathematically define this noise. Although a solution using non-linear differential solutions exist, work is still underway to create a generalized model which is simpler to implement and as accurate as the Split Step Fourier Transform (SSFT) method. As BVTs provide flexibility for granular tuning of modulation formats and bandwidths of their associated optical channels, the noise power experienced by each channel on other channels and on itself also becomes difficult to characterize. In principle, the NLIN of a Channel Under Test (CUT) can be mathematically defined as:

$$P_{NLIN} = \int_{\frac{-\delta\nu}{2}}^{\frac{\delta\nu}{2}} G_{NLIN}(f)df \tag{2.3}$$

where P_{NLI} is the noise power, $G_{NLIN}(f)$ is the Power Spectral Density (PSD) of the NLIN at a given frequency f. The two components of P_{NLI} are Self Channel Interference (SCI) and Cross Channel Interference (XCI). P_{SCI} is the noise power of the NLIN caused by the CUT on itself, whereas XCI is the noise power of the NLIN caused by the neighbouring channels on the CUT.

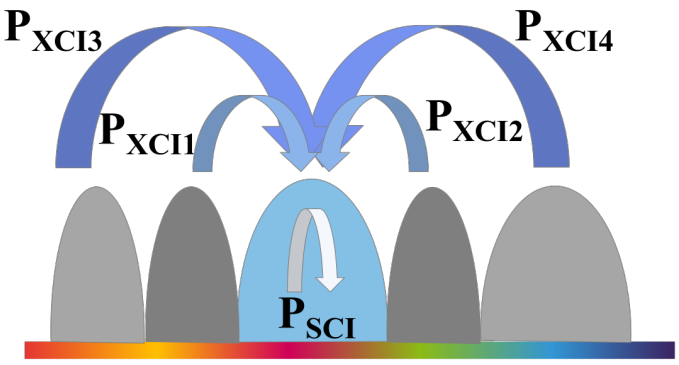


Figure 2.1: NLIN on the CUT due to four neighbouring channels and due itself.

Fig. 2.1 shows a flex-grid span with five optical channels. If we consider the central channel as the CUT, there are a total of five noise components, out of which four are XCIand one is SCI. In reality, the characterization and subsequent calculation are not as simplistic

as this example. Interested readers can find further details on NLIN using the GN model in Appendix A.

2.1.2 Optical Reach of Bandwidth Variable Transceivers

Flexible BVTs [3] can have varied settings for channel launch powers (-2 to 2 dBm), modulation formats (Quadrature Phase Shift Keying (QPSK) - 64Quadrature Amplitude Modulation (QAM)), and symbol rates (32 GBaud - 128 GBaud). This combination of channel power, modulation format, and symbol rate is known as a configuration. Several configurations can be created out of these parameters and a BVT can be tuned to this configuration, provided the optical channel traversing over the link has a GSNR that is greater than the BVT's minimum receiver OSNR for the given configuration. As an example, in Fig. 2.2, we show the minimum receiver OSNR for each of the 40 configurations available for a model BVT [12]. These values have been derived using analysis conducted by Teipen et. al [35]. The BVTs have different modulation formats, i.e., QPSK, 8QAM, 16QAM, 32QAM, or 64QAM, which allows the encoding of 2, 3, 4, 5, or 6 bits per symbol respectively. As the number of bits encoded per symbol increases, the data rate capacity of the BVT also increases.

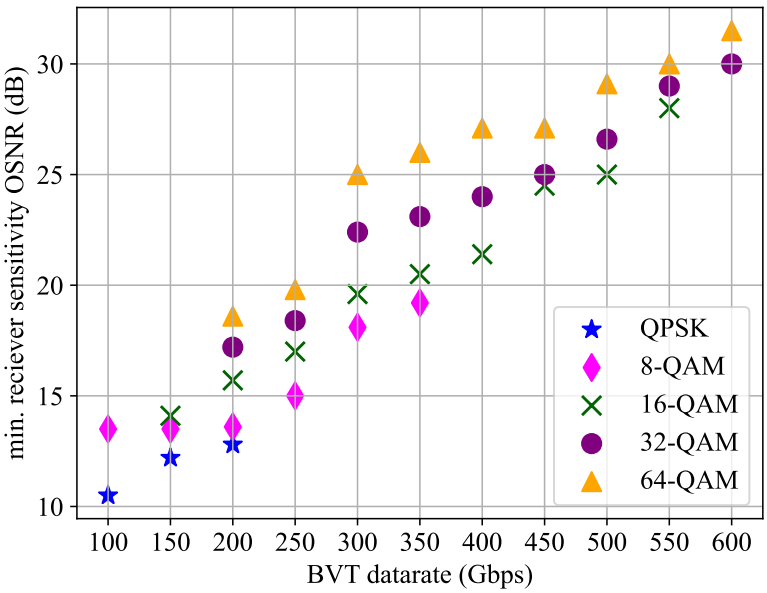


Figure 2.2: Data Rate (Gbps) vs minimum required OSNR (dB) for different modulation formats of a BVT. We assume perfect ILA amplification and homogeneous spans while neglecting ROADM losses.

[3] Usually termed as Flex-BVTs, however, in this dissertation, we use BVTs for brevity.

However, with the increase in modulation format and data rate, the minimum receiver sensitivity OSNR also increases. This is defined as the minimum OSNR needed at the receiving BVT to ensure error-free transmission of a transparent optical channel. As the distance traversed by the optical channel increases, higher data rates and modulation formats cannot be achieved. Therefore, one of the major tasks in network planning is to select an appropriate configuration for each BVT in the network in order to meet the requested traffic.

2.2 Network Planning Concepts

In the subsequent sections, we provide an introduction to long-haul optical network planning as well as define the terms which will be used extensively in the subsequent chapters.

2.2.1 Network Planning Terminologies

A national long-haul OTNs, as shown in Fig. 2.3, connects multiple cities in a country over long distances using optical fibers which are installed either underground or submarine. Since submarine optical fiber networks have different physical layer constraints, we restrict network studies to long-haul country-wide or continental networks.

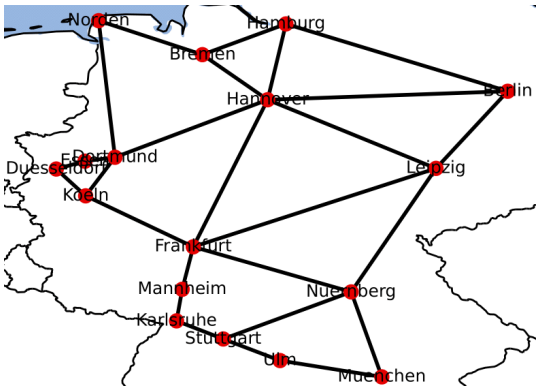


Figure 2.3: Example of the Nobel-Germany OTN, consisting of PoPs and links.

Specifically, network topologies of European countries are of major interest, due to the added advantage of dark fiber availability. Before discussing these topologies, we define certain terms used in physical-layer network planning. We begin by defining the optical network topology as a mathematical graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{D})$. The constituent parts of this graph can be defined as:

1. $\mathcal{V}$ - Set of PoPs in the network, where traffic can be added or dropped.
2. $\mathcal{E}$ - Set of links, where each link $L \in \mathcal{E}$ is defined as a collection of standard single-mode fiber spans. Each span $s \in L$ often varies in length to the other spans in the link owing

to geographic realities, and can connect either a pair of ILAs or an ILA and a node $N \in \mathcal{V}$.

3. $\mathcal{D}$ - Set of demands in the network which are the aggregated traffic capacity requests between two PoPs in the network. Each demand $d \in \mathcal{D}$ can be further split into multiple **lightpaths**, which are defined as a single optical channel of a BVT pair, having its own configuration, routes through the network, and a fixed central channel frequency on all the links it traverses through.

2.2.2 Optical Network Topologies

To conduct long-term network planning studies, multiple physical layer topologies need to be generated with fiber span information. These topologies are first retrieved from the Internet Topology Zoo [36] and then converted into a physical layer fiber and span topology, which is available as an open-source dataset [12]. Table 2.1 shows some of the topologies used in this dissertation[4].

Network	**Topology Parameters**			**Lightpath Lengths [km]**		
	Nodes	Links	Demands	Min	Avg.	Max
Nobel-Germany	17	26	121	34.50	558.47	982.00
Spain	16	27	135	112.79	871.81	1598.67
Sweden	25	29	286	20.00	1066.18	3533.34

Table 2.1: Topology and route statistics for European OTNs [12].

Although network topologies are rather easily available, for a physical layer aware network planning, there is a need to have the fiber topological information, specifically, the type and length of fibers in the networks, the noise figure and gain of each ILA in the network, and the add/drop and pass-through losses of all ROADMs in the network, located at PoPs. Such information is usually not available in the public domain. Therefore, several realistic assumptions need to be made for the underlying fiber map. Predominantly, the span lengths need to be heterogeneous in nature but always distributed around a normal value of 80 km. The noise figure and gain of ILA can be set according to the length of the span ingress into the ILA. Finally, minimal ROADM add/drop and pass-through losses need to be estimated, based on the band of operation.

Irrespective of the topology used in this dissertation, we design all multi-period network planning studies for 10 planning periods. Each planning period can be considered to be equivalent to a financial year, and planning for 10 years allows OTN operators sufficient time

[4]Interested readers can find topology statistics of other networks used in this dissertation (e.g. Nobel-EU and Abilene US) on the GitHub page [12].

to leverage their investments and achieve profitability. The fiber topology of these networks assumes a single SSMF fiber pair between the two nodes. For multi-band transmission, we consider different values for fiber attenuation, amplifier noise, launch power, and minimum required GSNR for each band, taking values from Correia et. al [37].

Once the fiber topology is established, a network traffic demand model is required to add demands over the planning years. Such traffic models are not available and need to be derived and corroborated by using several open-source information, like country-wide network traffic statistics from Internet Exchange Points (IXPs) and Data Centers (DCs) and other research works. The fiber topology, traffic model, and optical networking components serve as input to a network planning study. The output of such a planning study includes, but is not limited to, the overall achievable throughput of the network in each planning year, the number of BVTs used, and an overall cost estimate over the complete planning period.

2.2.3 Routing, Configuration and Spectrum Allocation

At the core of a flex-grid network planning study lies the RCSA algorithm. The RCSA can be defined as a list of tasks that needed to be performed, in order to place lightpaths in the OTN which are capable of carrying either the complete or a part of the requested traffic demand between the PoPs. We describe each of the tasks of the RCSA algorithm as follows:

1. Routing - For a given demand, a set of shortest paths needs to be calculated between the source and destination PoPs. Based on the requirements, the paths can be simple shortest paths, node-disjoint, or link-disjoint paths. Since the performance of BVTs decreases with an increase in distance, link length is usually chosen as the routing metric in these algorithms. Examples include Dijkstra's shortest path, Yen's k-shortest paths, etc.

2. Configuration Selection - Based on the results of routing, each path of each demand can have a set of configurations that would have their GSNR higher than the minimum receiver sensitivity. The configuration selection algorithm not only chooses the set of valid configurations but also finds the minimal number of candidate lightpaths that need to be deployed in order to meet the demand.

3. Spectrum Allocation - Each candidate lightpath in the OTN needs to satisfy the spectrum contiguity and continuity constraint. Contiguity refers to the constraint that a lightpath can only be assigned adjacent frequency slots on the flex-grid, whereas continuity refers to the constraint where the lightpath is assigned the same frequency slots on all optical links along the path. Different allocation strategies like first-fit, last-fit, or center-fit can

be used to place lightpaths efficiently. However, the first-fit allocation strategy is the most widely used solution in network planning applications.

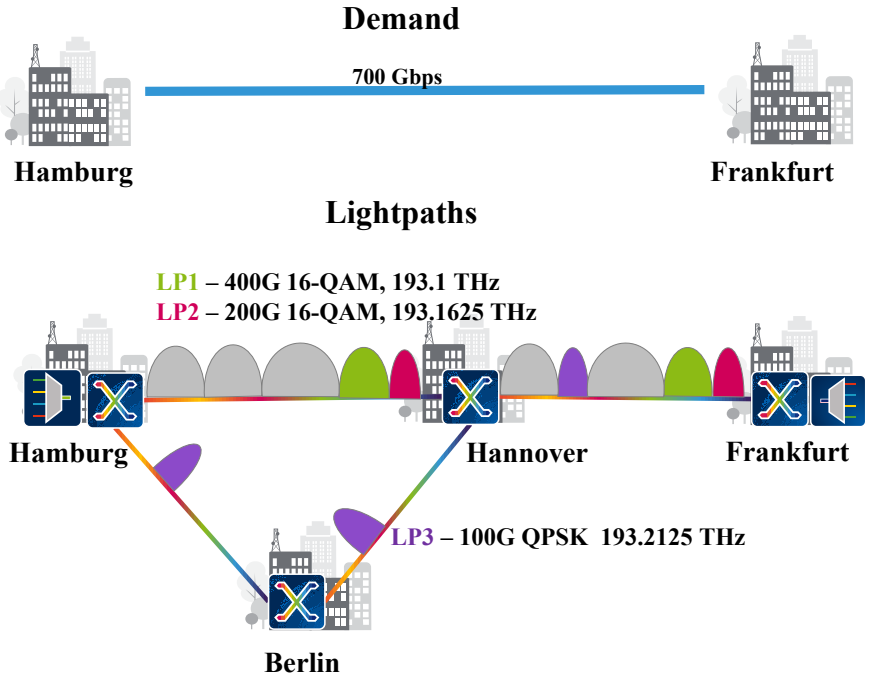


Figure 2.4: Simple RCSA for one demand between two PoPs in the Nobel-Germany network.

As a toy example of RCSA, we consider a demand in the Nobel-Germany network, as shown in Fig. 2.4. Let us assume that there is a requirement of 700 Gbps capacity between Hamburg and Frankfurt. First, we find two shortest simple paths between the two PoPs. These are Hamburg-Hannover-Frankfurt and Hamburg-Berlin-Hannover-Frankfurt respectively. Now, we need to find a set of configurations and the number of candidate lightpaths needed to satisfy the 700 Gbps. Let us assume, that due to a filled spectrum as well as reach constraints on the Hamburg-Hannover link (grey lightpaths), we can only place two candidate lightpaths, i.e., LP1 and LP2 on the first shortest path. LP1 carries 400 Gbps and is tuned to the central channel frequency of 193.1 THz on both Hamburg-Hannover and Hannover-Frankfurt links. Similarly, LP2 carries 200 Gbps of capacity and is tuned to the next available central channel frequency of 193.1625 THz. The remaining 100 Gbps capacity is satisfied by the third lightpath carrying 100 Gbps of traffic, routed on the second shortest path and tuned to 193.2125 THz of central channel frequency (the first available central channel frequency on all the links of the second route). Even though the previous central channel frequencies are available on Hamburg-Berlin and Berlin-Hannover optical links, the first central channel frequency which satisfies the wavelength continuity constraint is 193.2125 THz, due to a

spectrally full Hannover-Frankfurt link. In this way, a simple RCSA can be done for one demand.

In practice, RCSA for large networks is more complex and involves the use of both heuristics as well as meta-heuristics, as discussed in greater detail in Chapter 3.

2.3 Traffic Demand Models

As discussed in the previous section, demand is defined as the aggregate traffic capacity requests between two PoPs in an OTN. These traffic capacity requests (usually in Tbps) are classified as confidential data by long-haul OTN operators. Therefore, to create these traffic demands for network planning, researchers have relied on empirical traffic models like the gravity model [38]. The gravity model assumes traffic to be proportional to the human population of the source and destination. However, in present times, most of the traffic between two locations is exchanged between either DCs or IXPs and there exists a higher correlation of offered traffic to the node degree as compared to the human population. For example, Frankfurt and Duesseldorf are less populated cities as compared to Hamburg and Munich, but exchange traffic at higher rates [39].

Although network capacity studies have traditionally assumed Poisson arrival of light-path requests to achieve progressive loading, these scenarios rarely give actionable inputs to network operators, since realistically, network traffic exchanged is skewed in favour of well-connected nodes with higher population [40, 41]. Due to the lack of open-source network data, it is difficult to make assumptions about the required traffic in each network; however, some intelligent guesses can be made. Particularly, machine learning has also been shown to be used for traffic prediction in optical networks, although this is a specific solution for short to mid-term fluctuations in IPoWDM networks [42].

2.3.1 Initial Traffic Estimate

Furthering the mathematical definition of nodes, links, and demands as $\mathcal{V}$, $\mathcal{E}$, and $\mathcal{D}$ respectively, each node $\kappa \in \mathcal{V}$ has an associated metric on the number of co-located DCs (DC_κ) and IXPs (IXP_κ). For the network $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{D})$, we also define $\bar{N}$ as the average node degree of the network, and for each demand, $d_{i,j} \in \mathcal{D}$, $N = n_i + n_j$ represents the combined node degree of the source node i and destination node j.

To derive a model for the initial traffic estimate, we first assume a hierarchical structure of the co-located DCs and IXPs at each node location. Here, we make two assumptions, namely, *(i)* intra-node traffic (i.e. DC to IXP within the same location) is not accounted for in

a long-haul OTN traffic model, and *(ii)* inter-node traffic consists of only DC-DC and IXP-IXP interconnects. This modelling helps in reducing long-haul transparent network demands and allows intra-node aggregation networks to serve inter-node traffic.

Firstly, we define the number of IXP and DC connections to be established between source node i and destination node j in the network as,

$$\Delta_{i,j} = IXP_i \cdot IXP_j + DC_i \cdot DC_j \tag{2.4}$$

where IXP_k and DC_k are the number of IXPs and DCs in a given node $k \in \mathcal{V}$. This ensures we cover all the IXP-IXP and DC-DC connections between the nodes. However, there is also the need to remove any DC-IXP interconnects in line with the second assumption. Therefore the number of connections to be established can be modified as

$$\Delta_{i,j} = IXP_i \cdot IXP_j + DC_i \cdot DC_j - DC_i \cdot IXP_j - DC_j \cdot IXP_i \tag{2.5}$$

Simplifying the equation, we obtain:

$$\Delta_{i,j} = |DC_i - IXP_i| \cdot |DC_j - IXP_j| \tag{2.6}$$

From the combined node degree N, one egress link on the source node and one ingress link on the destination node are used for the traffic flow. Therefore, we multiply Eq. 2.6 with a combination equation of choosing two degrees from N. For sparsely connected nodes, the combination equation is replaced with the combined node degree. Hence, the total flow $\theta_{i,j}$ from node i to j is defined as:

$$\theta_{i,j} = \begin{cases} \binom{N}{2} \cdot \Delta_{i,j} & \text{if } N > 2 \cdot \bar{N} \\ N \cdot \Delta_{i,j} & otherwise \end{cases} \tag{2.7}$$

To estimate the initial traffic at planning period $t = 0$ between source i and destination j, the number of flows $\theta_{i,j}$ need to be multiplied by the per-interconnect demand in Gbps. For this, we chose two interconnect demands at the beginning of the planning period, namely, α_1 = 10 Gbps and α_2 = 7.5 Gbps. α_1 is assigned for flows between densely connected nodes, whereas α_2 is assigned for flows between sparsely connected nodes. Moreover, a constant β is set to 100 Gbps, which is assumed to be the minimum demand request between the source and destination, so that if $\Delta_{i,j}$ becomes nullified, there should be at least a 100 Gbps demand available. Therefore, the initial traffic at $t = 0$ for a demand $d_{i,j} \in \mathcal{D}$ is given as:

$$\tau(i,j,0)[Gbps] = \begin{cases} \alpha_1 \cdot \theta_{i,j} + \beta & \text{if } N > 2 \cdot \bar{N} \\ \alpha_2 \cdot \theta_{i,j} + \beta & otherwise \end{cases} \tag{2.8}$$

2.3.2 Traffic Growth Model

Once the initial traffic estimate is available, as a first approach, each demand requesting traffic can be simplistically assigned a single growth value. For this simplistic traffic growth model, we assume two growth parameters, namely, **expected growth** and **unexpected growth**. The values for expected growth are taken from available traffic estimates [43]. For each planning period $t > 0$, the offered traffic between nodes is defined as:

$$\tau_t = \gamma_t \cdot \tau(i, j, 0) \tag{2.9}$$

where γ_t is the expected growth in offered traffic [44]. For unexpected traffic growth, we modify γ_t with an increase of 30-90% between 2023 and 2026, bringing about an overall realistic increase of ca. 40% in the aggregate offered traffic growth for the given network topology. Since this growth does not capture the disaggregated growth of the traffic demands, a detailed empirical-based model is needed to capture the unequal growth effects. We now introduce a traffic model which decouples the traffic growth from the initial traffic demand.

Since the yearly traffic growth rates of around 25% are a common phenomenon [45], we define the overall expected traffic increase δ_t in a given planning period t as,

$$\delta_t = \tau_{t-1} \cdot \mathcal{X}_\tau \tag{2.10}$$

where τ_{t-1} is the aggregate traffic distribution of the previous planning period and $\mathcal{X}_\tau \leftarrow \mathcal{N}(\mu_\tau, \sigma_\tau^2)$ is random variable drawn from a normal distribution with μ_τ set to 0.25 and σ_τ set to 0.1. The mean of 25% growth is assumed from global network traffic growth predictions [43], whereas the deviation is extrapolated using the measurements reported by Soule *et. al* [46].

Further, we define the increase in traffic of individual demands, acknowledging the non-uniform traffic growth for each node in the network. For this, we first define a node-importance metric $\mathcal{I}_v$ for a node $v \in \mathcal{V}$ as,

$$\mathcal{I}_v = \Lambda_v + 100 \cdot (IXP_v + DC_v) \cdot C_v \tag{2.11}$$

where Λ_v and C_v is the eigenvector centrality and the closeness centrality metric of the node $v \in \mathcal{V}$ in the graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{D})$ respectively. IXP_v and DC_v are the number of co-located IXPs and DCs at node v.

From Eq. 2.11, it is inferred that nodes that are better connected and have a higher number of IXPs or DCs are the ones that will have higher importance. Once the importance of each node is calculated, we find the normalized value of this importance metric $\mathcal{I}_{v,norm}$, such that

the value for each node lies between 0 and 1, and the sum of all the normalized importance metric is 1.

$$\mathcal{I}_{v,norm} = \frac{\mathcal{I}_v}{\sum_{v \in \mathcal{V}} \mathcal{I}_v} \tag{2.12}$$

Fig. 2.5 shows as an example the normalized importance metric of each node in the Nobel-German network [30]. Nodes like Frankfurt have higher importance since they are both well connected and have a higher number of DCs and IXPs, whereas nodes like Munich, although having more number of DCs and IXPs, lie lower in the importance list, due to their location in the graph.

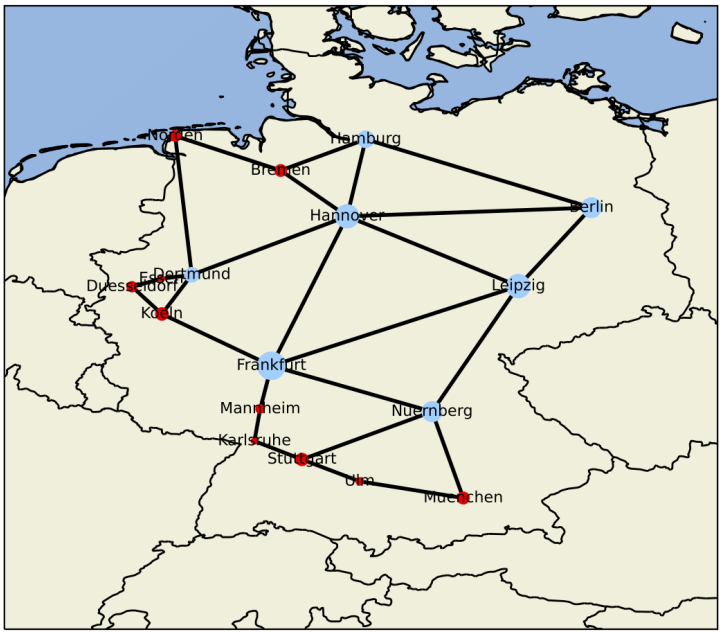


Figure 2.5: Nobel-Germany network, with node size proportional to the normalized node importance metric. Blue-colored nodes have a higher likelihood of exchanging traffic.

Since each demand $d \in \mathcal{D}$ has a source and a destination node $(i, j) \in \mathcal{V}$, a joint probability density function to select a demand using the normalized importance metric of each node can be calculated as,

$$\mathcal{F}_d(i, j) = \frac{\mathcal{I}_i \cdot \mathcal{I}_j}{\sum_{(a,b) \in \mathcal{V}} (\mathcal{I}_a \cdot \mathcal{I}_b)} \tag{2.13}$$

Therefore using Eq. 2.13, we draw a demand $d \leftarrow \mathcal{F}_d(i, j)$ and calculate the additional traffic request generated by it in the planning period as,

$$\delta(i, j, t)[Gbps] = D_{DC,t} \cdot (DC_{i,j}) \cdot e^{X_{DC} \cdot t} + D_{IXP,t} \cdot (IXP_{i,j}) \cdot e^{X_{IXP} \cdot t} \tag{2.14}$$

Here, $D_{DC,t}$ and $D_{IXP,t}$ are defined as the inter-DC and inter-IXP interconnection request at a given planning period t, and take the following values,

$$(D_{DC,t}, D_{IXP,t})[Gbps] = \begin{cases} (10, 40) & if\ t \leq 3 \\ (40, 100) & if\ 4 \leq t \leq 7 \\ (100, 400) & otherwise \end{cases} \tag{2.15}$$

$DC_{i,j} = DC_i + DC_j$, $IXP_{i,j} = IXP_i + IXP_j$, are the sum of the number of DCs and IXPs at the source and destination. $\mathcal{X}_{IXP} \leftarrow \mathcal{N}(\mu_{IXP}, \sigma^2_{IXP})$, and $\mathcal{X}_{DC} \leftarrow \mathcal{N}(\mu_{DC}, \sigma^2_{DC})$ are the IXP and DC growth rates drawn from their respective random normal distributions.

The values $(10, 40)$, $(40, 100)$, and $(100, 400)$ refer to the additional traffic request base values in Gbps generated between each DC and IXP at the source and destination respectively. These values are assumed to increased every four planning periods, due to increase in technology adoption, leading to interconnect traffic demand increase. Moreover, the demand request also depends on an exponential increase from the base values of DC and IXP interconnects. Specifically, these growth rates are drawn from a random variable with a Gaussian distribution, whose mean values are 0.35, 0.25 and standard deviation values are 0.1, 0.05 respectively [43, 46].

This process continues till the sum of all additional traffic from all drawn demands reaches the aggregate additional traffic δ_t as defined in Eq. 2.10. Therefore, using this method, we can achieve a skewed demand distribution, with more traffic flowing between nodes of importance. This is especially helpful in multi-period network planning studies, where the demands need to not only increase, but also introduce stochastic events in the simulation.

2.3.3 Traffic Model Validation

Since our proposed traffic model is an important input to the network planning study, we need to compare and validate the model with the traditional gravity-based traffic model [47], and with the CISCO VNI report on traffic growth for three network topologies under test.

Network	Gravity Model scaling parameter (k)	Total Peak Traffic (2021)[Tbps]	Assumed Market Penetration [%]	Peak Traffic per operator (2021) [Tbps]	CAGR [%]
Nobel-Germany	100	127	25	31.75	26
Spain	10000	62	33	20.66	23
Sweden	100	22	100	22	24

Table 2.2: Traffic modelling parameters for the Gravity Model [47] and CISCO VNI forecast [43].

The gravity model defines the traffic $t_{s,d}$ flowing between a source s and destination d in a long-haul OTN as $t_{s,d} = k \cdot \frac{P_s \cdot P_d}{L^{\phi}}$, where P_s and P_d are the population metric at source and destination and L is the distance between them. k and ϕ are the scaling factors for the traffic. For this comparison, we assume the population metric P_s and P_d to be the sum of

IXPs and DCs at the source and destination respectively. ϕ is set to 2, assuming two different types of traffic are being carried. Scaling factor k varies according to the network and is chosen using trial and error to achieve similar order of magnitude as the initial aggregate traffic mentioned in the 2021 CISCO forecast highlights [45]. Another assumption made is that the market penetration is uniformly distributed amongst the long-haul network operators in the country. The traffic increase rates for the gravity model and the forecast are the same as the Compounded Annual Growth Rate (CAGR) mentioned for each of the countries in the CISCO forecast. The discussed parameters are assimilated into Table 2.2.

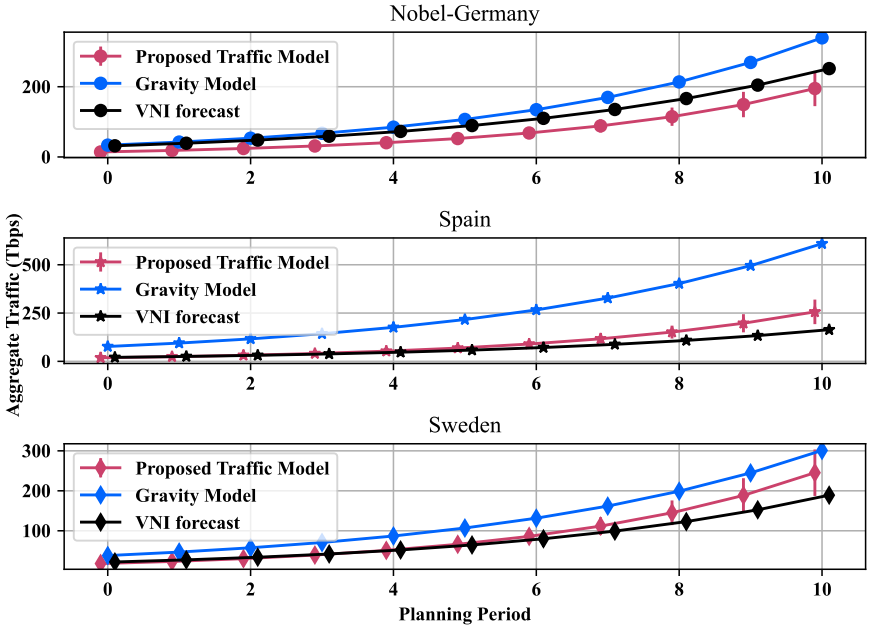


Figure 2.6: Traffic Model comparison for the Nobel-Germany, Spain, and Sweden OTNs.

As seen in Fig. 2.6, the proposed traffic model follows the traffic growth forecast closely. The gravity model, on the other hand, tends to deviate from the forecasted growth, despite having the same CAGR. It can be argued that the k and ϕ parameters of the gravity model can be further optimized to mimic the VNI forecast. However, the proposed traffic model does not require any parameter optimization and is generic enough to be used for long-haul networks holding different graph properties. Finally, independent of this validation study, the results of the traffic model for the Spanish network show a similar trend to Fig.5 in Lopez *et. al.* [48], justifying the accuracy of this traffic model.

2.4 Disaggregated Optical Network Device Control

As defined in Chapter 1, SDN seeks to separate the control plane from the data plane of a given networking device. This means that all the control decisions which earlier occurred on the device itself, are now offloaded to a centralized controller, which communicates with the networking device using various interfaces. To ensure automation in OTN, three main components are needed. These are data modelling languages, protocols, and standards. Each of these components is discussed in detail.

2.4.1 Data Modelling Languages

Data modelling languages are high-level languages that define the topology, connection, and operation of OTN devices. Since devices have many parameters which are layered or inter-dependent, such languages are required to be logically structured, preferably in a tree format. Another requirement is the ability of the data model to be converted into any data format like Extended Modelling Language (XML) or Java Single Object Notation (JSON), which allows for information transport using APIs. The data models explored and used in this dissertation are:

1. Structured Management Information (SMI) - Based on one of the earliest data models, i.e. ASN.1, SMI is the data model used by SNMP. SMI is divided into module definitions, object definitions, and traps. Module definitions describe the behaviour of an information module, whereas object definitions describe the syntax of a managed object. Finally, traps are used to create notifications to transmit a device's behaviour information.

2. YANG - Defined by Internet Engineering Task Force (IETF)'s RFC 6020, YANG are the preferred data models for modern SDN communication protocols [49]. Each data model in YANG is known as a module and consists of nodes and a leaf, where a node is an independent variable within the module, and each leaf is an attribute of the node it is attached to. Related nodes can be grouped together in the same container, and lists are used to store nodes in a particular sequence. The biggest advantage of YANG is its ability to import nodes defined in other modules.

Based on the ownership of the data models, they can be classified either as open models or native models. Open models are either written and maintained by standardization bodies, or by open-source consortiums and are designed to be OTN device independent. Conversely, native models are defined by OTN device vendors and are usually proprietary to the vendors.

Although native models have the advantage of modelling specific behaviour, open models are required for future interoperability and device disaggregation.

2.4.2 Protocols

A protocol provides a method for the SDN controller to communicate with the OTN devices. Protocols are built in conjunction with the data model and ensure that the modules and nodes of the data models are encoded into data-interchange formats like XML or JSON and can be configured by controllers using pre-defined Remote Procedural Calls (RPCs). In this dissertation, we delve into two protocols, namely NETCONF and RESTCONF.

NETCONF, as defined by the IETF, is a protocol used to install, manipulate, and delete configurations on network devices [50]. NETCONF allows for the usage of RPCs and encodes data into XML format for communication. Using a pre-defined port, a NETCONF client (usually the controller), can remotely connect to an OTN device hosting the NETCONF server and send XML-based commands for various device-based operations. The request and response XML need to be mapped to the correct YANG model at both the server and client sides. In this dissertation, we use NETCONF for south-bound changes to the network, i.e., communication from the SDN controller to the OTN devices.

RESTCONF refers to the usage of RESTful applications via HTTP-based APIs. Using RESTCONF, controller or OTN device data can be manipulated via Create, Read, Update, Delete (CRUD) actions. For RESTCONF, data is encoded in JSON. For our work, RESTCONF-based requests are used for higher layer applications like network planners and orchestrators, based on APIs as defined by the SDN controller being used.

2.4.3 Standards and Open-Source Consortium

For SDN-based multi-hierarchical communication, the decision of the combination of data models, encoding formats, and protocols, along with the detailed definition of each of the modules, RPCs and APIs is taken by standardization bodies like International Telecommunication Union Telecommunications Standardization Sector (ITU-T) or IETF. However, the process of standardization is slow and can take many years before an industry-wide standard is released. However, the requirement and business use-case to disaggregate devices in OTNs needs quicker solutions. Interoperability reaps benefits for network operators since they can choose different devices from different OTN vendors and create a streamlined solution based on the traffic demands. However, a pre-cursor to this is that the network operators must support the same data models and protocols, which are required by the SDN controller planned to be used. Therefore, open-source consortiums like OpenROADM, OpenDaylight, ONF, and

OpenConfig aim to create methodologies so as to pave the way for standardization. In this dissertation, we focus on the OpenROADM consortium for data models and protocols.

The OpenROADM Multi-Source Agreement (MSA) is a multi-operator multi-vendor initiative, which aims to standardize the specification of OTN devices like ROADMs, BVTs, and ILAs [51]. OpenROADM provides YANG models for device control as well as network and service management. Broadly, the three models are as follows:

1. Device model - OpenROADM device model describes the collection of YANG modules needed to describe OTN devices. These are usually bounded to south-bound RPCs and are used to synchronize the device information to and from OpenROADM compliant OTN devices.

2. Network model - OpenROADM network model provides YANG models to describe the topological components of the network at different abstraction levels. These abstractions allow the controller to focus on network-related functionalities like routing, and path computation elements.

3. Service model - These YANG models are needed to create, delete, or modify services and describe the north-bound interfaces to orchestrators. Operations related to optical service management on different layers like the transport and data layer are also dealt with in this model.

2.4.4 Open-Source SDN Controller

OpenDaylight is an open-source project maintained by the Linux Foundation, and over the years has established itself as a stable SDN controller. OpenDaylight controller works on a microservice-based architecture, which allows the users to enable and disable required applications before launching the controller. This versatility allows for the development of various applications on top of the OpenDaylight codebase. One such application is Transport-PCE, which is designed specifically to manage OTN infrastructure. TransportPCE supports YANG models from OpenROADM and exposes NETCONF RPCs to connect south-bound to OTN devices, as well as north-bound REST APIs which can be consumed by higher layer applications.

Fig. 2.7 shows an architectural view of the TransportPCE controller with its constituent components. The OpenROADM device, network, and service models are mapped to certain internal components which interact with them. Each of the internal components are also coloured according to their functions. Red-colored cylinders and red arrows portray databases and data flows respectively. Blue-colored rectangles represent internal microservices which

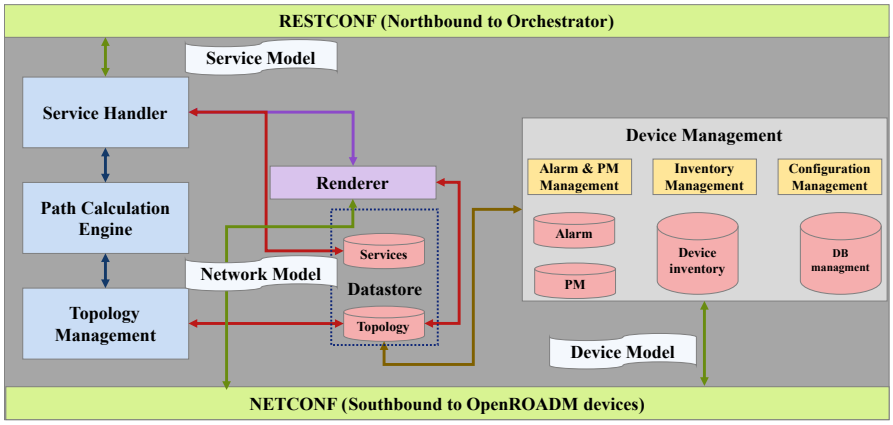


Figure 2.7: Components of the TransportPCE controller used for OpenROADM model interaction [24].

are invoked for specialized functions. Yellow colored rectangles represent management functions available for device management. Finally, green colored rectangles and green arrows are the implemented interfaces and the different OpenROADM models communicating using these interfaces. The major components and their functions are further listed as follows:

1. Service Handler - Receives requests from higher layer applications via north-bound APIs and checks with the Path Computation Element (PCE), if a path for the required service is possible. If yes, the service is created and added to the running service list.

2. PCE - This module takes care of the RCSA and tries to place the optical service on to the available spectrum on the links. It also takes into account several network constraints such as node exclusion, shared risk group exclusion, and maximum latency.

3. Topology Management - This component builds the topology using the network model of OpenROADM and also notifies any changes in topology to the Service Handler and the PCE.

4. Renderer - Converts the path description provided by the service handler into actual changes to be made on each OTN device along the path. If there are no interfaces available on those devices, the Renderer creates a request to add the missing interfaces. Following the OpenROADM service model, two service paths, one from source to destination and other from destination to source are created.

5. Device Management - Device management component consists of several databases and business logic which configures the devices, manages its inventory and manages the device alarms and performance monitoring data. These databases are updated using OpenROADM device models.

In this dissertation, we use the combination of OpenROADM with OpenDaylight TransportPCE as the SDN controller in a multi-domain open-source service management Proof-of-Concept, details of which are discussed in Chapter 5.

2.5 QKD Schemes

In this section, we discuss the salient features of a widely used QKD protocol which is implemented in many of the commercially available QKD devices [29, 52]. Due to its low quantum bit-error rate values at longer SSMF lengths, BB84 is also selected as the QKD transmission protocol in the QKDN trusted node deployment studies in Chapter 6.

2.5.1 BB84 Protocol

BB84 protocol was proposed by Bennett and Brassard in 1984 [53]. In this protocol, Alice encodes the key information onto single photonic qubits using an electro-optical modulator. BB84 requires two bases to encode information, namely a computational basis (binary 0 or 1) and a diagonal basis (non-orthogonal polarization angle θ). Alice uses a hardware-based random number generator to first choose one of the bases and then using a uniform probability distribution, chooses the binary bits 0 or 1. Therefore, each qubit transmitted by Alice can have one of the four possible states of 0, 1, $0_{\pi/4}$, and $1_{\pi/4}$, with an equal probability of 0.25.

Variable	Range	Unit
Mean photon number	0.1-0.3	-
Fiber attenuation SSMF	0.2-0.22	dB/km
Additional loss	1-3	dB
APD quantum efficiency	10-20	%
Quantum channel repetition rate	1-1.25	GHz

Table 2.3: Range of input parameters for secure key generation rate using BB84 protocol [54].

At the receiver side, Bob detects the incoming qubits and randomly selects a measurement basis for each photon. To do this, a passive 50/50 beam-splitter partitions the incomeing qubits into two different polarization analysers (one for 0 and 1, and the other for $0_{\pi/4}$ and $1_{\pi/4}$). If Bob measures the incoming qubit with the correct basis, he achieves a complete information retrieval, otherwise, his result is 100% uncorrelated with the information encoded by Alice. During the sifting process, when Alice and Bob receive each other's measurement bases, they correlate them with their own bases and discard the non-matching bits. If it is assumed that Eve has no access to the trusted nodes, the key exchanged in this process is deemed information-theoretic secure.

2.5.2 Secure Key generation rates for BB84

In QKD communication, the SKR between an Alice-Bob QKD pair is the rate at which Alice and Bob can exchange information and generate symmetric keys at their respective location [55].

For the BB84 protocol, the SKR is defined by [54]:

$$SKR = \mu \cdot L_f \cdot L_o \cdot P_d \cdot L_p \cdot \psi \tag{2.16}$$

where μ is the mean photon number, L_f is the attenuation of the SSMF, L_o is the additional losses related to bending or coupling, P_d is the quantum detection efficiency of the Avalanche Photo Diode (APD), and ψ is the quantum channel transmission rate.

These variables operate within a particular range and depend on various factors, like the attenuation constant of the SSMF spans, the distance between Alice and Bob, or the efficiency of the APD used in the QKD devices. Table 2.3 suggests an acceptable range for these parameters, which could affect the SKR.

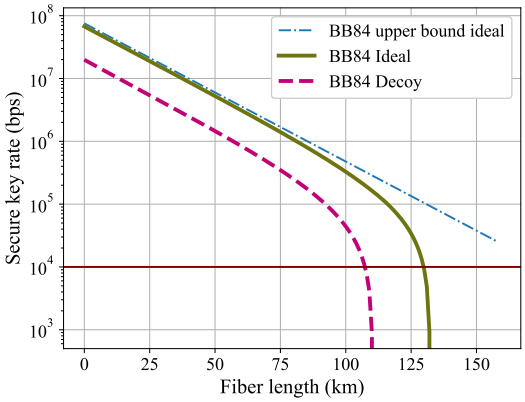


Figure 2.8: Comparison of SKR vs SSMF length.

To improve the robustness of the photon being transmitted by the BB84 protocol, decoy states can be introduced. In a decoy-state extension of BB84, Alice is required to transmit a decoy pulse of qubits which does not contain any useful information. Since Alice can now additionally randomly choose to send a "normal" state or a decoy-state, Eve cannot distinguish between the two types of states. Moreover, Alice and Bob can use the decoy state for sifting and privacy amplification, thereby leading to an almost linear decrease of the SKR as the fiber length increases [56].

We further compare the performance of BB84 *ideal* (single photon source) state with the BB84 *decoy* state. Using Eq. 2.16, we can find the SKR in bits per second.

Fig. 2.8 shows the reduction in the SKR as compared to the distance between Alice and Bob. The attenuation is assumed to be 0.22 dB/km, the other parameters are selected as the lower-bond values given in Table 2.3. We see that although the decoy state does not perform as well as the ideal state, it is still able to provide a high SKR around 80 km, which is the mean length of SSMF in OTNs. Therefore, the BB84 decoy-state protocol is a prime candidate for long-haul QKDN deployment.

Chapter 3

Optical Transport Network Planning Algorithms

In this chapter, RCSA algorithms for physical layer planning of long-haul OTN networks are discussed and the research question **RQ1** is answered.

The contribution can be divided into three distinct parts. In the first part, a heuristic to maximize the configuration of each BVT for a given planning period is proposed [3]. Then, using an ILP, this solution is extended to a multi-period OTN planning scenario [4]. Finally, an RCSA solution using regenerators is presented and the advantages of using BVTs in a flex-grid translucent OTN is studied [5].

This chapter is structured as follows; for each of the contribution needed to answer the first research question (**RQ1**), an introduction and motivation is offered, followed by the state-of-the-art for the particular contribution[1]. Then, methodologies and results are discussed, and finally, the chapter is concluded.

3.1 Heuristic for Configuration Selection in Optical Networks (HeCSON)

With advances in BVTs supporting software tunable channel configurations [43, 57], network planners possess a new degree of freedom to select the best **configuration**, i.e. combination of data rate, modulation format, and Forward Error Correction (FEC) for each lightpath. In-house network planning tools and open-source software require planners to manually enter this information [58]. However, when multiple combinations of modulation formats (QPSK, 8QAM, 8/16QAM, 32QAM, 64QAM), data rate (100-600 Gbps) and FEC (0%,15%,27%)

[1] Although the state-of-the-art has significantly improved since these contributions, this dissertation restricts the discussions to the state-of-the-art available at the time of the contribution

are possible, manual optimization becomes tedious. Since NLIN in a fiber depends on path parameters like neighbouring channels, span lengths, etc., choosing configurations for each lightpath affects the QoT of the neighbouring channels of those lightpaths, leading to a race-around degradation condition in the network. To undertake network feasibility studies, knowledge of the best possible data rate for a source-destination pair is required. Hence, there is a need for an algorithm that quickly provides a list of viable configurations for each demand, considering the calculated GSNR of each lightpath.

3.1.1 State-of-the-art Analysis

For a QoT metric to measure the performance of each channel, typically GSNR is used. The definition of GSNR follows from Eq. 2.1. To model P_{NLI}, the Full Form Enhanced Gaussian Noise Model (FF-EGN) [59], has emerged as a more accurate model as compared to the other GN models [60]. However, FF-EGN is calculation intensive and does not scale well in terms of time and resources for large networks with many traffic demands. To cater to this, [61] presents a Accurate Closed Form Enhanced Gaussian Noise Model (ACF-EGN) model, which shows good results in end-to-end links, but has not been verified on heterogeneous network topologies[2].

As discussed in Section 2.1, the reach and capacity limitations of each BVT depend on the GSNR along the route taken by the lightpath. More importantly, an accurate calculation of NLIN needs to be included in planning algorithms. In recent years, GN models have been developed in order to provide an accurate estimation of NLIN to the system under study. The initial models, however, were either not closed-form or had a high error in the modeled NLIN, as compared to the SSFT method. This could potentially lead to optimistic network deployment plans and inaccurate estimations of the OTN's capabilities [59]. Recently, research has begun on creating accurate closed-form models, either using mathematical derivation and assumptions or by using Machine Learning on the various noise parameters [62]. For a network planning application, a trade-off between GSNR accuracy versus path calculation time is of utmost importance since the RCSA algorithms use the physical layer calculation model several times to generate the list of candidate lightpaths. A solution with very low NLIN computation time, as well as reasonable accuracy[3], is therefore, the most desirable solution candidate.

To address this trade-off between accuracy and calculation time, we present *HeCSON*, which is a heuristic to maximize the data rate of each lightpath in the network, while ensuring that the QoT of each lightpath is not violated.

[2]Refer to Appendix A for a brief overview of GN models
[3]$< 1dB$ deviation from SSFT method

3.1.2 HeCSON and Path Calculation Tests

The proposed heuristic, *HeCSON*, can be categorized as an RCSA algorithm for a single planning period. Here, the routing, configuration selection, and spectrum assignment are handled as independent sub-problems and solved one after the other, for each demand $d \in \mathcal{D}$, where $\mathcal{D}$ is a set of source-destination pair bidirectional demands in a given network.

The networks used in the simulation of the heuristic are Germany50 (50 nodes, 88 links, 662 demands), Nobel-EU (28 nodes, 41 links, 378 demands), and Norway (27 nodes, 50 links, 351 dem nds) [30]. Since there is no open-source physical network information available for these networks, we have modeled the length of spans and type of amplifiers from distributions of real networks [12].

In order to achieve “new" configurations like 600G 64QAM with 15% FEC [63], we start by creating various combinations of data rates between 100-600 Gbps, in steps of 50 Gbps, bits per symbol from 2 to 6, and FEC of 0%, 15% or 27%. From these, we select the configurations with a symbol rate between 32-72 GBaud, which leaves us with 60 possible configurations. The worst-case transceiver performance, given by the minimum required GSNR, is available for some configurations in [43]. We extrapolate these and assign a minimum required GSNR to each configuration, which is then used as a selection and validation metric in *HeCSON*.

The aim of this heuristic is to find an explainable compromise between the computation time and accuracy of path calculation engines, by using different GN models.

3.1.2.1 HeCSON Workflow

Using Fig. 3.1, we explain the heuristic. Starting with the list of demands and a list of all possible configurations, we calculate Dijkstra’s shortest path (Dijkstra-SP) for all the demands and then perform Pre-Selection by calculating the linear GSNR for all demands based on the demand’s path length. The configurations for which the linear SNR drops below the minimum required GSNR are removed from the available configuration list of a particular demand. The configuration list is then ordered by first maximizing data rate and then minimizing channel bandwidth amongst the ordered maximized data rates.

The first available configuration for each demand is then taken from the list and is placed using a first fit channel allocation. Once all demands are allocated an initial configuration, we iterate over each demand to check if it is validated and begin Configuration Selection. After selecting the next non-validated demand, ACF-EGN block calculates the total ACF-EGN GSNR and checks if it is higher than the minimum required GSNR of the assigned configuration [4]. In case it is, the configuration is assigned to the demand, otherwise, we

[4] low GSNR with dash-dot lines

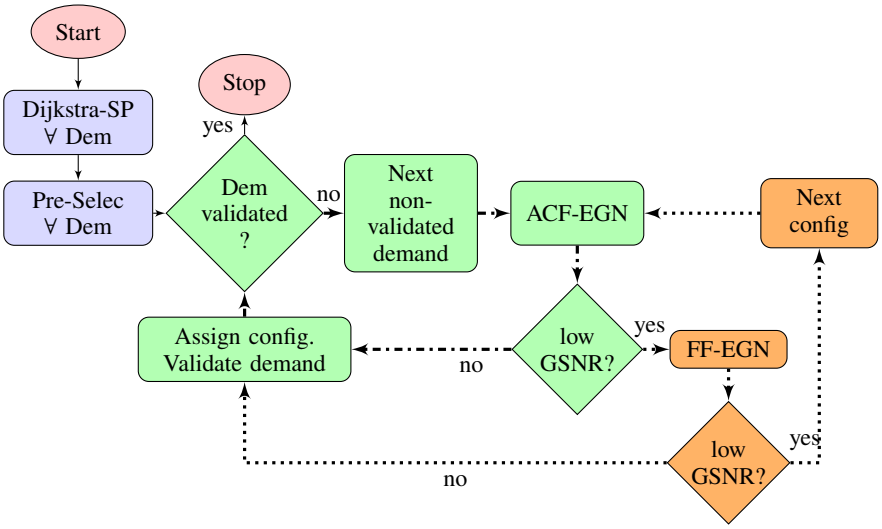


Figure 3.1: HeCSON's workflow.

execute Configuration Validation[5]. For the current configuration, the total FF-EGN GSNR is calculated (FF-EGN block). If that too falls below the minimum required GSNR [6], the configuration is removed from the channel allocation matrix and the next configuration available for the demand is placed (Next config). Otherwise, in case of validation, we assign the configuration to the demand and mark it as "validated" for *HeCSON*. The demands for which no configuration is selected due to not satisfying the minimum GSNR requirements are blocked.

3.1.2.2 Path calculation tests

We compare *HeCSON* with a configuration selection based on a pure ACF-EGN model and a pure FF-EGN model. In both "pure" cases, every time a placed configuration's total GSNR falls below the minimum required total GSNR, the next configuration is placed and the total GSNR is re-calculated. The two major parameters we evaluate in our study are the total network throughput, defined as the sum of all the placed demand data rates, and the total execution time, defined as the time taken for every case to run using a Java program on an Intel Core i7 processor with 16 GB RAM. The other results of interest are the number of blocked demands, which are the demands blocked either due to lack of spectrum slots or suitable configuration; and spectral efficiency, defined as the ratio of overall network throughput (Tbps) to the total occupied spectrum on all links in the network (THz). It should be noted that we consider throughput results only for those demands which are placed by all three cases. We look at the throughput to show that the use of multi-configuration, flexible-

[5]blocks connected with dotted lines in Fig. 3.1
[6]low GSNR connected with dotted lines

grid transceivers in the network provide higher data rates for demands. Execution time is important for future online planning applications so that the algorithm can do real-time GSNR recalculation for all neighboring channels whenever a new demand is added.

3.1.3 Results and Discussions

Upon running *HeCSON* for the Germany50 network, we see that only 22 of the 60 possible configurations have been used. This is shown in Fig. 3.2, where a majority of the demands are placed between 200 and 400 Gbps configurations. We see a substantial number of demands selecting 300 and 400 Gbps 15% FEC based configuration, which provide higher spectral efficiency compared to 27% FEC based configurations. The application of our pre-filtering step, as explained in Section 3.1.2.1, helps in reducing the execution time vastly and removing configurations for which even the linear SNR would fall below the minimum required GSNR.

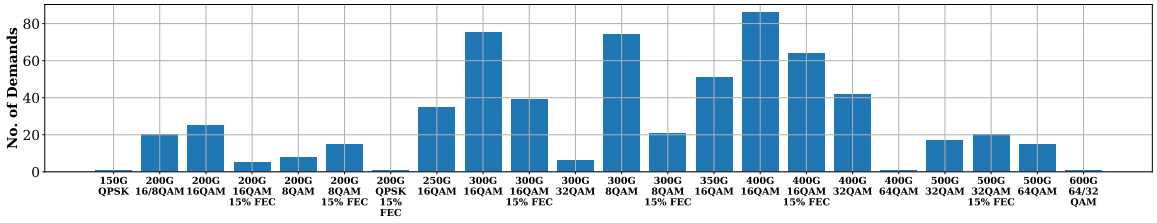


Figure 3.2: Distribution of demands for each of the 22 selected configurations out of the 60 possible onses in the Germany50 network.

As shown in Fig. 3.3 (a), HeCSON provides a total network throughput of 199.85 Tbps, which is comparable to 209.20 Tbps suggested by the FF-EGN case.

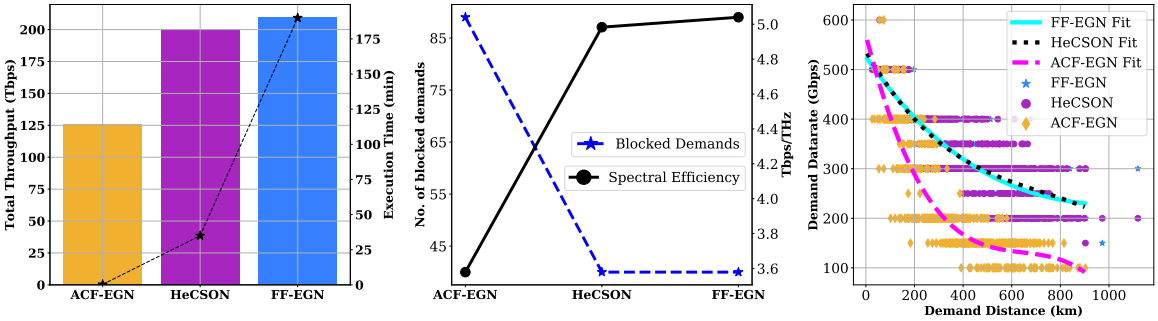


Figure 3.3: For the Germany50 network, (a) network throughput and execution time, (b) number of blocked demands and spectral efficiency, and (c) polynomial fit of data rate vs demand distance for all three cases.

HeCSON completes execution in 35 min (star plot in Fig. 3.3 (a)), compared to 190 min for FF-EGN case. Reduction in time is helpful for planners to test out different scenarios

and configurations quickly. Comparing with ACF-EGN case, the overall network throughput is increased by more than 75 Tbps. For further generalization of our analysis, we also look at the spectral efficiency (Fig. 3.3 (b)) for each of the cases. We see that *HeCSON* has a spectral efficiency of 5.02 Tbps/THz, comparable with FF-EGN case's spectral efficiency of 5.04 Tbps/THz and having a 60% increase as compared to ACF-EGN. For all three

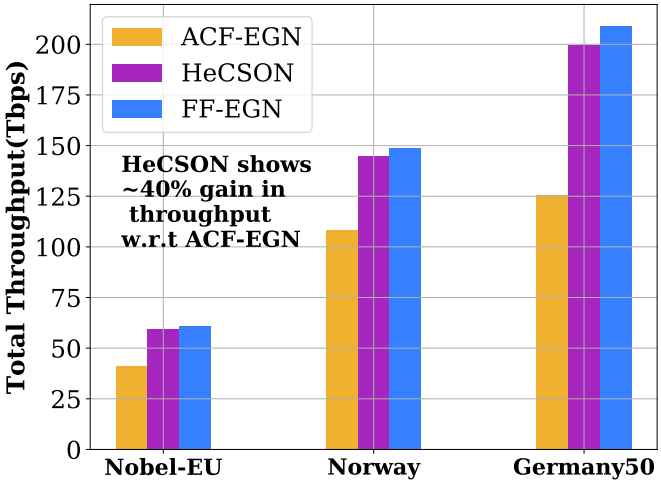


Figure 3.4: Total throughput for three networks.

cases, at least 40 demands are blocked in the Pre-Selection stage, due to a lack of adequate spectrum slots on some links. Additional 49 demands are blocked in ACF-EGN case due to not satisfying the minimum required GSNR. This is because ACF-EGN, like other approximations mentioned in [59], overestimates P_{NLI}. Looking at the chosen data rate for each demand, we see a clear decrease in data rate, as the distance increases. When fitted using a third-order polynomial, we see that *HeCSON*'s fitted curve closely follows FF-EGN's curve. Running the path calculation tests on different-sized networks, we see from Fig. 3.4 that the throughput follows similar trends in each of them. For the three networks, namely, Nobel-EU, Norway, and Germany50, we consider maximum possible bi-directional demands between non-adjacent source-destination pairs, i.e., 378, 351, and 662 respectively. Compared to ACF-EGN, *HeCSON* provides a throughput gain of approximately 40%. We can therefore conclude that *HeCSON* is network agnostic and can be used in different optical networks for transceiver configuration assignment.

3.1.4 Summary and Conclusions

To summarize, we propose *HeCSON* as a new heuristic for configuration selection in heterogeneous optical networks. About 40% gain in the total network throughput as compared to

the conservative pure ACF-EGN case and an 87% decrease in execution time as compared to the accurate pure FF-EGN case is obtained. This heuristic serves as a starting point to further develop optimized solutions for configuration selection in BVTs.

3.2 Multi-period Planning with HeCSON (MP-HeCSON)

As long-haul OTN operators realize the increased demand for high-speed internet at home and work, many fixed-grid OTNs need to be upgraded to meet growing demands. These demands have shown a big increase in 2020 due to the recent COVID-19 pandemic, where many network operators reported a 30-50% increase in broadband traffic in Q1 2020, compared to Q4 2019 [64]. Given the worldwide impact on logistics, operators have started to invest in software configurable BVTs, which offers opportunities to increase network throughput in order to carry the growth in offered traffic. Offered traffic is defined as the total throughput of the network if all demands are met by adding sufficient BVTs in the network [7].

To mimic the growth of traffic over multiple planning years, the initial traffic estimate model discussed in Eq. 2.8 (Section 2.3.1) is used. For the traffic growth over the planning years, two traffic growth models *viz.*, **Expected** and **Unexpected** are introduced, and shown in Fig. 3.5 for a ten-year planning period. The expected traffic growth model follows CISCO VNI CAGR estimate, whereas the Unexpected growth model induces an increase of more than 40% in the second and third planning year, simulating black swan events which might increase the internet traffic substantially [45].

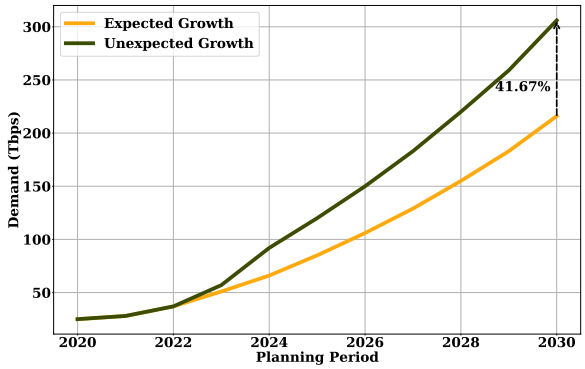


Figure 3.5: Aggregate Offered Traffic Growth models.

In this work, a multi-period extension of the HeCSON heuristic, called ***MP-HeCSON*** is proposed. *MP-HeCSON* offers two types of network upgrade schemes, namely *Scheme 1* and *Scheme 2*. Briefly, *Scheme 1* invokes the combinatorial heuristics of *MP-HeCSON* to first upgrade the in-operation lightpaths before adding new BVTs every planning period. *Scheme 2*

[7] see Section 2.3

on the other hand does not disrupt in-operation lightpaths and only adds new BVTs into the network to meet traffic requests.

3.2.1 State-of-the-art Analysis

Traditionally, lightpath configuration in network planning consists of algorithms exploiting Integer Linear Programming (ILP) as well as heuristics. A link state heuristic for optical networks [65] is of particular interest; however, no insights on its usability in a multi-period scenario are provided. The SNAP Algorithm [66] offers progressive traffic loading, randomly allocating traffic between nodes without considering realistic growth. Both works also assume homogeneous spans across their physical topology, limitations of which have been discussed in our previous work [3].

Once the path calculation model is selected, network planners need to look at algorithms to place candidate lightpaths into the available spectrum. As discussed in Section 3.1, the additional requirement of configuration selection makes RCSA an NP-hard problem. Several works have looked into RCSA from a flex-grid network planning viewpoint. Broadly, RCSA can be solved by using an ILP [67, 68], or a joint solution using combinatorial optimization [69, 70]. The first solution relies on modelling the RCSA problem mathematically. Although it is possible to model and add additional objectives or constraints, the computational efforts are huge due to the large solution space and the need to model non-overlapping channel allocations for all candidate lightpaths deployed across the network. Also, such a solution is difficult to scale for large network studies. The second solution involves splitting the RCSA problem into independent smaller modules and solving them separately [71]. Although optimality is not guaranteed, such algorithms can converge to near-optimal solutions faster, thereby allowing planners to invoke multiple scenarios.

3.2.2 MP-HeCSON and Planning Study

The multi-period planning algorithm *MP-HeCSON* is shown in Fig. 3.6. It takes as input the topology, along with information on the number of DCs and IXPs at each source-destination pair (i, j). The initial offered traffic using Eq. 2.8 is calculated and an initial Routing and Spectrum Allocation (RSA) using a weighted probabilistic routing based on Yen's k-Shortest Path Algorithm and the number of continuous empty frequency slots in each of the paths is performed. This heuristic for routing adds randomness in choosing the candidate path list for each (i, j). For spectrum assignment, a first-fit channel allocation strategy is used.

The objective of the algorithm is to minimize the number of lightpaths added to the network for planning period t while trying to meet the offered aggregate traffic. Readers

may note that one lightpath can be associated with one and only one BVT in the scenarios discussed. To begin, we take as input the candidate path list of all demands ($\mathcal{D}$), all currently provisioned lightpaths in the network (L) and the traffic matrix (τ) calculated using Eq. 2.8[8] and Fig. 3.5.

For each candidate path $d_{i,j,t} \in D_t$, we compute a list of lightpaths $L_{i,j,t} \subset L_t$ already present between nodes (i, j) and additional offered traffic $\theta_{i,j,t}$, defined as $\tau_{i,j,t} - \tau_{i,j,t-1}$. Using the 60 different channel configurations for each BVT are generated. The list of valid channel configurations $C_{i,j,t}$ is then filtered using *HeCSON* [3]. For every provisioned lightpath ($\lambda \in L_{i,j,t}$), a set of channel configurations $\rho_{i,j,t} \subset C_{i,j,t}$ is found, having its channel bandwidth BW_ρ less than or equal to the provisioned lightpath's channel bandwidth BW_λ.

This ensures the upgrade of provisioned lightpaths at the same central channel frequency, without additional bandwidth usage. Iterating over $\rho_{i,j,t}$, which is sorted according to the highest data rate, the first ρ which has a higher data rate than λ is selected. If such a ρ is found, λ is updated with ρ and the spare capacity generated can be used to satisfy some of the additional traffic in the future planning periods, thereby reducing $\theta_{i,j,t}$.

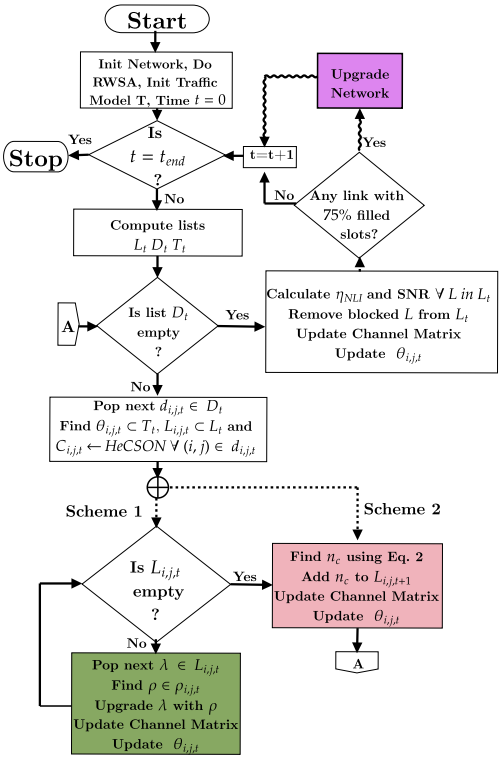


Figure 3.6: MP-HeCSON flow in a multi-period scenario with the proposed schemes.

After the provisioned lightpaths are upgraded, the number of additional lightpaths x_c of each valid channel configuration c are found. These lightpaths are needed to carry the

[8]Ref. Section 2.3

remaining $\theta_{i,j,t}$ using the ILP shown in Eq. 3.1. This ILP minimizes integer x_c, such that the configured total data rate can only be over-provisioned by δ Gbps. For the network studies in this work, δ is set to 100 Gbps allowing for an overprovisioning of 100 Gbps for every demand in each planning period. A power independent NLIN constraint, η_{NLI}, is also defined. This constraint follows definitions and calculations of ACF-EGN model [62]. The NLIN constraint restricts the total η_{NLI} being added every t to the sum of provisioned lightpaths η_{NLI}. The lightpaths are then added to $L_{i,j,t+1}$, defined as a set of configured lightpaths between nodes (i, j) for the next planning period.

$$\text{minimize} \sum_{c \in C_{i,j,t}} x_c, \tag{3.1}$$

$$\text{subject to: } \theta_{i,j,t} \leq \sum x_c \cdot DR_c < \theta_{i,j,t} + \delta, \tag{3.2}$$

$$0 < \sum x_c \cdot \eta_{NLI_c} \leq \sum_{\forall \lambda \in L_{i,j,t}} \eta_{NLI_\lambda} \tag{3.3}$$

After additions are completed, we also check on the number of empty frequency slots in the link. If more than 75% of the frequency slots have been filled up in any of the given links, it warns operators about upcoming frequency slot saturation in the configured fiber pair so that they plan a localized physical upgrade by exploring either additional bands or utilizing available dark fiber pairs [66].

Network	Nodes/Links/Demands	Span variation (km)	Noise Figure (dB) [66]	Node Deg. (min/avg/max)
Nobel-Germany	17/26/272	30-120	4.3	2/3.05/6
US Abilene (US12)	12/15/132	20-100	4.3	1/2.5/4

Table 3.1: Reference topology information [12].

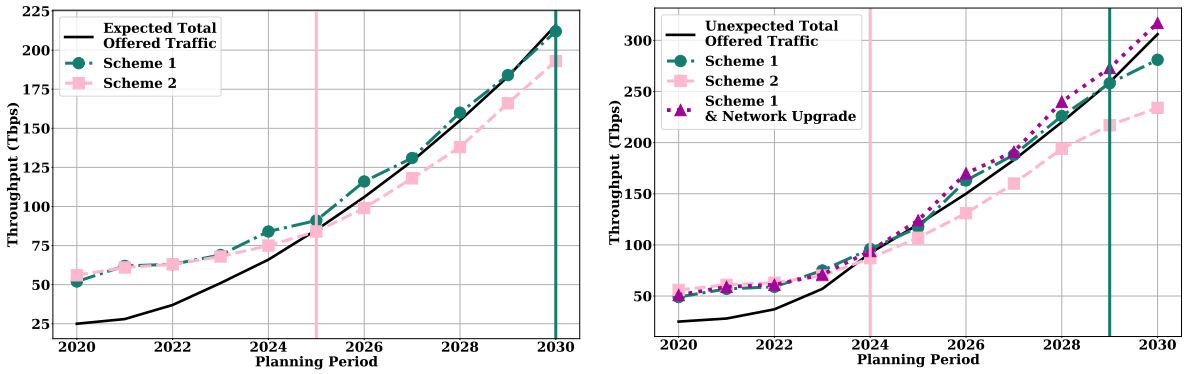
Using the discussed flow, a multi-period network study for the 17-node Nobel-Germany OTN [12]. Single fiber pair C-band operation is assumed and fiber spans end into variable gain EDFAs having constant noise figure as shown in Tab. 3.1.

3.2.3 Results and Discussion

Scheme 1 and *Scheme 2* are evaluated in the aforementioned traffic scenarios. It must be noted that an additional scheme, where all BVTs are placed according to their maximum possible data rate has been previously evaluated in Section 3.1, compared to which, both *Scheme 1* and *Scheme 2* provide a higher network throughput. For expected traffic growth shown in Fig. 3.7a, we see that offered traffic can be satisfied by both schemes till 2025. After which, we observe that the overall expected throughput for *Scheme 2*[9] cannot meet the offered traffic. The cause may be attributed to the lack of additional frequency slots on some links, which

[9]dashed line with squares in Fig. 3.7

cause additional lightpaths (x_c) found in Eq. 3.1 to remain unprovisioned. This translates to a loss of approximately 40 Tbps expected throughput at 2030. In the same scenario, *Scheme 1* meets offered traffic without the need of a network upgrade for the entire planning period t.



(a) Expected traffic throughput

(b) Unexpected traffic throughput

Figure 3.7: Yearly aggregate throughput of Nobel-Germany topology from 2020-2030.

Similarly, for unexpected traffic growth (shown in Fig. 3.7b), upgrading provisioned lightpaths using *Scheme 1* enables the operators to utilize the C-band for five additional planning periods. However, *Scheme 1* also shows a downward trend in provisioned lightpaths, beginning from 2029. This infers that network operators would have to upgrade their network with new fibers or equipment, in order to meet the traffic growth. The question then arises as to when must the operator upgrade their network. In the evaluated topology, it is observed that three links cross the occupied slot threshold of 75% in 2027, which is when additional capacity must be planned for. Post upgrade (either fiber doubling or band doubling), the offered traffic can be met by the algorithm.

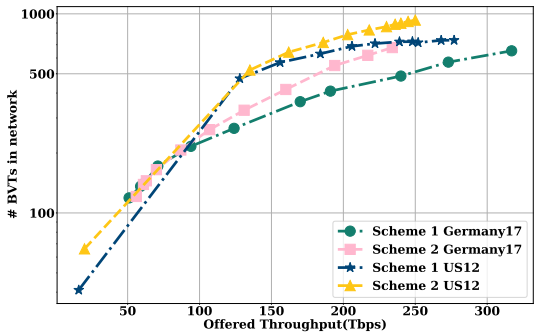


Figure 3.8: BVTs vs Throughput for Nobel-Germany and US Abilene.

Fig. 3.8 shows that *Scheme 1* achieves a 40% increase in throughput for the same number of BVTs in Nobel-Germany topology. Conversely, similar throughput can be achieved by *Scheme 1* utilizing 25% lesser BVTs. For the 12 node US Abilene topology (US12 in

Fig. 3.8)[10], a 30% increase in throughput by using 18% lesser BVTs is observed. The reduction may be because of long-distance demands in the US12 topology, due to which offered traffic can only be met using more BVTs at lower modulation formats. However, *Scheme 1* still provides savings on the number of BVTs.

3.2.4 Summary and Conclusions

With increased opportunities for remote work and meetings, long-haul OTN traffic is bound to grow at a rate higher than initial forecasts. Pragmatically, operators may use BVT over-provisioning to meet such surges in traffic. The proposed solution *MP-HeCSON*, evaluated with two network planning schemes, shows a 40% overall increase in aggregate offered traffic for a realistic Nobel-Germany OTN. It is also shown that the offered traffic can be carried up to five planning periods, by efficiently planning the data rate and bandwidth usage of configured BVTs. For steeper and unexpected traffic growth, an estimate on when to plan a physical network upgrade is also provided.

3.3 Regenerator-based RCSA

The availability of BVTs offers network planners to flexibly allocate resources without excessively overprovisioning the network [18]. However, this advantage increases the complexity of planning algorithms, which must cater to varying data rate, channel bandwidth, modulation format, and corresponding receiver sensitivity for each configuration of the BVT. Furthermore, the allocation of flex-grid channels, with a minimum spacing of 12.5 GHz, requires the calculation of NLIN every time channels are added, removed or reconfigured in the network [3].

Network planning for flex-grid OTN undertakes several inter-dependent steps [72], starting with, a) routing each Traffic Demand (TD) in the network, b) assigning a set of transceivers to the TD and assigning each transceiver with a configuration, such that the traffic request is met, and c) assigning central channel frequencies (or wavelengths) to each transceiver in the network while maintaining wavelength continuity for every routed TD. Additionally, Regenerator Points (RPs) maybe placed to extend optical reach on longer routed TDs. In a BVT enabled flex-grid long haul optical network, adding RPs can also lead to reduction in the overall number of deployed BVTs, to meet the yearly traffic [73].

One of the major concerns of OTN operators is to cope with ever-increasing traffic. The fixed-grid C-band capacity has been increased with the use of flex-grid BVTs [18]. BVT

[10]US12 Topology has the same topological properties as US-Abilene Topology introduced in Table 3.1.

allows operators to select the best channel configuration for a given lightpath in terms of data rate, modulation format and FECs. Possible BVT configurations differ in terms of modulation formats, data rate and FEC. However, as demands keep increasing, alternatives to increase the network capacity furthermore have to be evaluated.

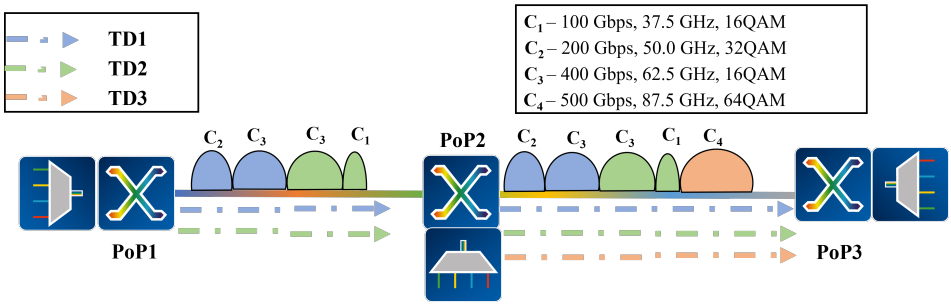


Figure 3.9: An example network with three nodes, where 4 BVT pairs are assigned to three different TDs.

Fig. 3.9 shows an example flex-grid point-to-point link with add/drop nodes PoP1, PoP2, and PoP3. Each of these adddrop nodes consists of BVTs, which can achieve multiple configurations: C_1 to C_4. Let us assume there exist two TDs between PoP1 and PoP3, each having a traffic request of 600 Gbps (TD1) and 500 Gbps (TD2), and one TD between PoP2 and PoP3 of 500 Gbps. Given the set of possible configurations, a configuration assignment algorithm may use the GSNR to calculate the physical reach limitations for each configuration. In Fig. 3.9, TD1 can be achieved by two BVTs, each configured to C_3 and C_4 respectively, as shown by the blue channels. Similarly, TD2 can be achieved by three BVTs (green channels). Finally, assuming that the QoT metric is satisfied, TD3 can be achieved with a single BVT configured to C_4.

Therefore, in this work, the advantages of using RPs along with BVTs is explored. A novel BVT allocation heuristic for the RCSA problem in regenerator-enabled networks is proposed. Applying a realistic traffic model on the Nobel-Germany network, it is seen that using RPs not only allows operators to meet the yearly requested traffic, but also leads to a reduction of up to 25% of the total number of BVTs, and savings of up to 50% in spectrum utilization, as compared to a solution without regeneration.

3.3.1 State-of-the-art Analysis

Since our solution provides novelty in the BVT allocation part of the larger RCSA workflow, we specifically survey three areas, namely, a) regenerator placement strategies, b) QoT calculation, and c) configuration selection for BVTs.

In flex-grid networks, regenerators are added by placing additional BVTs at intermediate nodes to achieve higher end-to-end data rates. However, these regenerators need to be

minimized to reduce additional costs incurred by OTN operators. Since regenerator placement is an NP-hard problem, a comprehensive solution using a genetic algorithm is available [74]. However, such algorithms have high computational and time complexity and simpler heuristics are often more desirable, which provide network planners with quicker insights. The authors in [75] use a gradual RP deployment heuristic along network planning periods, to minimize the cost of placing RPs. The solution uses transmission reaches of different BVT configurations, along with the requested demand throughput to place RPs in the network. However, higher data rates such as 400 Gbps, which could significantly reduce the number of RPs, are not considered.

Different BVT configurations have their own minimum receiver sensitivity GSNR. The calculated GSNR should be always greater than the minimum receiver sensitivity GSNR of the corresponding receiver. Previous works presented use an FF-EGN Model-based[11] QoT estimation to find the GSNR for each lightpath [3, 76]. However, these values can change as new lightpaths are added to the network, which may lead to re-calculating the GSNR for all lightpaths. To avoid such an increase in computation, some improved QoT calculation strategies can be exploited. The authors in [77] explore some pre-computing strategies for traditional RSA problems. A similar pre-computation of GSNR, in the first period of planning, provides simple bounds for subsequent configuration selection algorithms, while assuring that no lightpath falls below the minimum required GSNR value.

Several RCSA algorithms have been studied by the authors in [4, 68, 73] , each with different objective functions, based on certain assumptions. For example, Walkowiak *et.al*. assume that a bandwidth of only 37.5 GHz is requested every time, which reduces the complexity of their configuration selection [73]. However, in real-life deployment scenarios, the requested bandwidth that can be allocated to each BVT can vary between 25-100 GHz. This increase in generality also increases the solution space of BVT deployment algorithms.

Thus, in a multi-period network planning scenario, all the algorithms such as routing, RP allocation, GSNR calculation, BVT configuration generation, and spectrum assignment need to be used in a sequenced fashion to extract the network throughput and spectral efficiency. Algorithms such as routing and RP allocation can be applied only in the first planning period to reduce simulation time, the others should be executed in all periods due to traffic growth. The whole workflow of BVT configuration in multi-period planning is discussed in the next section.

[11] Ref. Appendix A

3.3.2 Regenerator-based BVT allocation

The BVT allocation workflow is divided into three major parts, namely, routing and regenerator placement (occurring once before the planning begins), GSNR calculation, and BVT allocation (occurring every planning period).

Let us assume that several TDs can share the same source-destination pair and each TD can be assigned multiple BVT pairs, one each at the source and destination, respectively. To extend the reach of longer routes, both lumped and distributed (Raman) EDFAs are considered by the algorithm when the GSNR is calculated. A pre-calculation of the GSNR with only linear ASE noise allows filtering out the configurations that do not guarantee the minimum GSNR, thereby avoiding unnecessary NLIN calculations. It is implied that for the same transmission path, the GSNR keeps degrading as more channels are added into the network [59]. The overall RCSA algorithm for all planning periods is shown in Fig. 3.10. The first step of this algorithm is to route all TDs using Dijkstra's shortest path. Routing needs to be performed in advance, not only to calculate the linear GSNR for each TD but also to find the TDs which would need an RP.

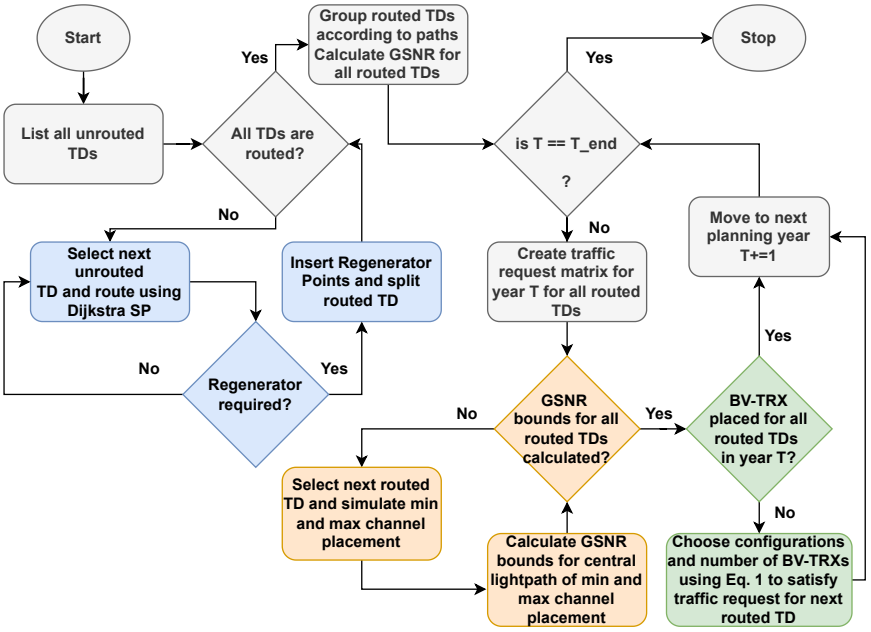


Figure 3.10: BVT allocation workflow diagram for all planning periods ©IEEE/IFIP 2021.

Once the path of each TD is found and the scenario of BVT with RPs is considered, the required RPs are placed according to two reach-limiting constraints, namely the maximum number of spans in the route and the total route length. Both constraints are determined based on the network characteristics. The algorithm traverses the links of the path, calculating the current distance from the start node and the number of spans. If the distance and/or the number of spans exceed the corresponding constraint, an RP is allocated to the node

between the current and the previous link. Then, the current distance and number of spans are reinitialized and the algorithm searches for the next location of RP until the end of the route is reached. This RP allows replacing the long TD by a trail of several shorter TDs. This continues for each TD, till there are no further violations of the reach-limiting parameters.

Calculating the GSNR is an important step to find the appropriate BVT configuration, however, the exact channel conditions in a multi-period traffic scenario are unknown. This is because the NLIN calculations need the channel placement information along the transmission path of the BVT. To avoid this race-around condition, all TDs sharing the same source, destination, and route are grouped together. By grouping TDs, all the requested data rate in each planning period can be summed into a "group data rate". Firstly, only those BVT configurations, whose minimum GSNR is greater than the GSNR of the route of the grouped TDs are filtered. Then two configurations with the minimum and maximum data rate are calculated. After this, two homogeneous channel based on the minimum and maximum configurations are simulated around the center of the C-band to meet the group data rate. For these two channel placements, the GSNR of the central channel is calculated based on the Enhanced Gaussian Noise Model (EGN) model [59]. Only the GSNR of the central channel is calculated, on the assumption that the central channel has the highest amount of XCI and therefore the worst-case GSNR value. Therefore, two worst-case GSNR values are obtained from the minimum and maximum lightpath assignment thereby creating a bound on the GSNR value for a particular group of TDs in the given planning period.

It is pertinent to mention that although this calculation step may lead to deviations from the actual GSNR values, the accuracy can be sacrificed for faster computation time. In the proposed method, the EGN model needs to be run only twice per grouped TD per planning period. Such a method can enable planners to quickly run multiple studies and plan different solutions for the network under study.

After completing the GSNR calculation (orange boxes in Fig. 3.10) for the current planning period, the BVT allocation process starts. This begins by selecting the traffic request matrix of the current planning period, for all the routed TD groups. Then, for each routed TD group, a list of applicable BVTs is created, whose GSNR satisfies the minimum OSNR, while keeping a small margin for future lightpath additions (1 dB in simulations undertaken in this work). The available BVT with the highest and the lowest data rate defines the lowest and highest number of lightpaths, respectively, that can be assigned to the TD.

For example, one 400 Gbps lightpath or four 100 Gbps lightpaths can be used for a grouped TD of 400 Gbps, assuming that BVTs are able to achieve either 100, 200, 300, or 400 Gbps on the given route. To find the exact configuration of BVT to be allocated to each grouped TD (for example, 2*100 + 1*200 Gbps), the optimization problem shown in Eq. 3.4 is used,

which minimizes the total bandwidth of BVTs assigned to the grouped TD, such that the sum of the data rates of all the lightpaths is at least the TD data rate in every period.

$$\text{minimize} \sum_{\forall i \in N} BW_i \cdot x_i \tag{3.4}$$

$$\text{subject to:} \sum_{\forall i \in N} DR_i \cdot x_i \leq DR_{TD}, \tag{3.5}$$

$$\text{subject to:} GSNR_{TD} \geq GSNR_{x_i} + M, \forall i \in N \tag{3.6}$$

where x_i is a binary variable assigned as 1 if a lightpath belongs to N. N is a set of lightpaths allocated to a grouped TD, BW_i, DR_i are the BVT bandwidth and data rate, respectively, DR_{TD} is the grouped TD data rate, $GSNR_{TD}$ is the lower bound of the GSNR values pre-calculated for a given grouped TD, $GSNR_{x_i}$ is the BVT minimum GSNR, and M is the additional GSNR margin.

The minimization is done by comparing the bandwidth of all BVT configurations. It is noted that the value of N is a variable, based on the multiple configurations which meet the TD group's data rate. These are built through the use of an algorithm searching combination with repetitions [78]. That is, all BVTs configurations assigned to the TD will have their minimum GSNR higher than the worst-case previously calculated GSNR. Then the configuration with the lowest spectral bandwidth will be selected as a solution.

In case the workflow is unable to place any BVT configuration for a routed TD, the TD is blocked. Afterward, optical add/drop equipment is configured [79] and the system performance is assessed [80].

3.3.3 Simulation Results and Discussions

For the comparison of total network throughput, the required number of BVTs, and occupied frequency range, we consider one planning period as one year and undertake ten-year planning (2021-2030) by creating a TD matrix for the Nobel-Germany network based on the traffic growth model discussed in Section 2.3. Fig. 3.11a shows the system throughput of our proposed heuristic (with and without RPs) and results of *MP-HeCSON* [4] with respect to the requested traffic. It can be observed that the solution with RPs is 15% closer to the requested overall traffic in the early planning years, as compared to the solution without RPs. Both solutions have a lower margin to the requested traffic as compared to [4] in the early years, as shown in Fig. 3.11a. However, in the later years, i.e., 2027 onwards, all tend to converge to the requested traffic.

As it is shown in Fig. 3.11b and Fig. 3.12 the total wavelength range occupied by all BVTs in a network is lower when RPs are used. Depending on topology and traffic (number of

nodes, span length, and requested system throughput) spectrum utilization is observed to be two times lower in a solution with RPs. Fig. 3.11b shows with respect to the left axis the occupied spectrum range in terms of the equivalent number of wavelengths required for a conventional fixed grid of 50 GHz. A higher number of BVTs along with a lower number of equivalent wavelengths in use compared to the algorithm presented in [4] at early planning years (as shown in Fig. 3.11b and Fig. 3.13) is explained by the usage of different wavelength allocation algorithms.

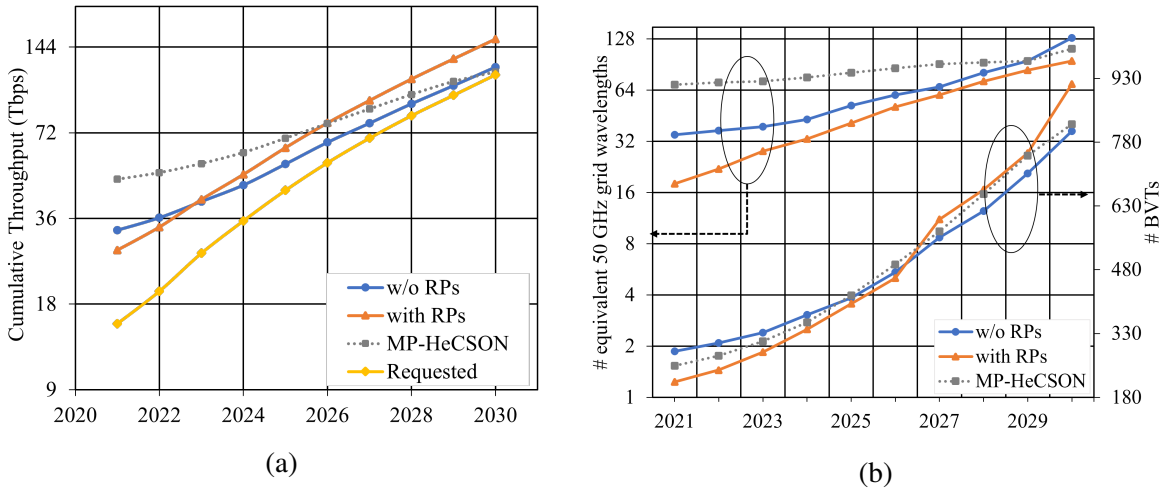


Figure 3.11: Simulation results and comparison with [4]. (a) system throughput, (b) number of BVTs and equivalent wavelength number calculated on a 50 GHz fixed frequency grid.

The relative difference of simulation results of both algorithms for the scenarios without RPs during the ten-year planning is shown in Fig. 3.12. The proposed algorithm generates BVT configurations that lead to effective spectrum utilization and system throughput in the first eight planning years. On the other hand, starting from 2026, the solutions are better in terms of the number of used BVTs.

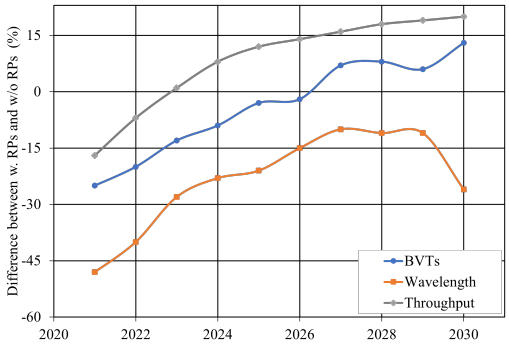


Figure 3.12: Relative growth of system throughput, number of BVTs, and wavelength number in configuration with RPs with respect to configuration without RPs.

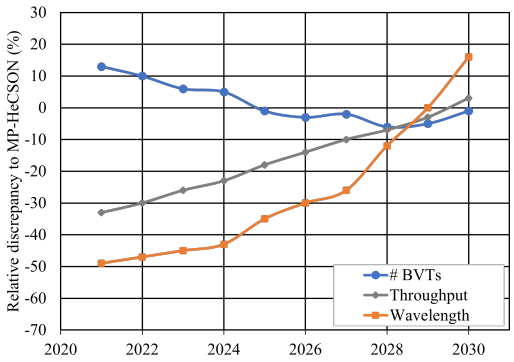


Figure 3.13: Comparison of relative change in throughput, number of BVTs, and wavelength number when RPs are used, with respect to MP-HeCSON [4].

Using the proposed algorithm, more than 96 wavelengths of 50 GHz equivalent are required in the last year of the ten-year planning without using RPs. This means that for the last year, a C+L-band network should be built assuming the absence of additional dark fibers. However, the usage of RPs in the workflow allows planners to keep all wavelengths within the C-band for the whole planning. As shown in Fig. 3.11b and Fig. 3.12 the number BVTs allocated by the algorithm is lower for some combinations of network topology and TD matrices when RPs are used. This shows that along with an improved spectrum utilization, the usage of RPs may reduce the number of required BVTs for certain combinations of network topology and TD matrix. This effect can be explained by the grouping of routed TDs after inserting RPs and before allocating BVTs, as it is shown in Fig. 3.10. Besides, the usage of RPs splits the long paths into shorter ones reducing also the OSNR degradation. Hence BVTs with higher data rate and more advanced modulation formats can be used.

3.3.4 Summary and Conclusions

A BVT allocation workflow, which allocates BVTs by accounting for linear impairments of the transmission line as well as for nonlinear interference penalties of neighboring optical channels is presented. The usage of regeneration points allows for improving spectrum utilization as well as the overall network throughput.

Chapter 4

Multi-band Network Planning Studies

In this chapter, research question **RQ2** is answered. As shown in simulation results of Chapter 3, C-band OTN systems are expected to saturate in the final planning periods and therefore are unable to meet the requested traffic. This leads to underprovisioning which translates into SLA violations or inability to offer services to new customers.

To avoid such a situation, operators need to seek solutions to upgrade their networks. These upgrades could either be in the form of lighting additional dark fiber pairs with a new C-band OLS per fiber pair or investing in new OLS equipment which allows for the use of additional bands (L-band and S-band) and exploit more throughput per deployed SSMF.

The contribution towards **RQ2** can be split into three sections. In Section 4.1, state-of-the-art RCSA solutions for the C-band are compared with an RCSA solution for C+L and S+C+L-band. In Section 4.2, the choice of terminal devices on multi-band upgrades is studied, and a simple CAPEX study is to compare different terminal architectures. Finally, Section 4.3 explores different methodologies to select band and fiber upgrade strategies in a multi-band scenario.

4.1 Multi-band RCSA

One of the major concerns of OTN operators is to cope with ever-increasing traffic. In Chapter 3, the C-band throughput has been shown to increase with the use of BVTs [18]. However, as demands keep increasing at an exponential rate, alternatives to increase the network capacity furthermore have to be evaluated.

In this work, different RCSA alternatives are studied to cope with the increase of the demands from an OTN operator's perspective. The network is defined as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{D})$, where $\mathcal{V}$ is the optical add/drop nodes and E the links consisting of SSMF pairs with heterogeneous span lengths, considering a finite discrete planning horizon $\mathcal{T}$ and a set of demands

$\mathcal{D}$ per year. Each demand $d \in \mathcal{D}$, identified by the source and destination, has a data rate DR_d that is configured at the beginning of the planning horizon following the initial traffic model proposed in Eq. 2.8 (Section 2.3). The traffic growth in other planning years follows the unexpected growth curves shown in Fig. 3.5. Each demand $d \in \mathcal{D}$ may be fulfilled by one or more lightpaths (LP_d) in the network, which may have different configurations and hence, data rates. The objective to achieve for each demand can then be defined as:

$$DR_d \leq \sum_{\forall lp \in LP_d} DR_{lp} \tag{4.1}$$

while reducing the over-provisioning as well as the number of BVTs. Moving on, we compare different alternatives that operators have in order to fulfill the data rate increase of the demands.

Further, in this contribution, the RCSA algorithms move from single-band to multi-band optical networks. First, a C+L-band capacity planning study is undertaken, followed by an S+C+L-band capacity planning study using BVTs[1]. Regenerator-based C-band approach with the band expansions are also studied and the advantages of multi-band network deployment in terms of throughput and underprovisioning are shown.

4.1.1 State-of-the-Art Analysis

Over the past decade, several transmission experiments have been undertaken to increase the traffic carried by a single fiber. Of these, two broad methods were, *i)* increasing the total number of frequency slots by exploring the S- and L-bands, or *ii)* increasing the modulation rate of a BVT, to achieve higher data rates per frequency slot.

Early works of exploring L- and S-bands were able to transmit 32 10 Gbps non-return to zero Amplitude Key Shifted (NRZ-ASK) signals over 160 km of Non-Zero Dispersion Shifted Fiber (NZDSF) across the S-, C-, and L-Bands [81]. With the introduction of QPSK and QAM signals, a record transmission of 160 DWDM channels, spread across the C- and L-bands over 240 km was shown [82]. Despite these seminal works, communication equipment manufacturers failed to add such capabilities to their product inventories, as the overall demand for such a solution was assumed to be low. A few years later, early works exploited advances in coherent detection using DSP to achieve record transmission of 200 Gbps Dual Polarized 16 Quadrature Amplitude Modulation (DP 16-QAM) signals, achieving 69.1 Tbps C+L-band transmission over low-loss pure silica core fibers [83].

Since operators are facing an exponential increase in traffic, network deployment studies focus on expanding to multiple optical transmission bands [37, 84]. With similar methodolo-

[1]The optical communication bands are introduced in Section 1.1.2

gies and tools, these studies look at maximizing the amount of traffic carried in each band, while focusing on power control and network upgrade strategies.

As an attempt to procure a solution for multi-band network planning in polynomial time, a combinatorial optimization-based approach can be implemented. This involves splitting the RCSA problem into independent smaller modules and solving them separately [71]. Without the guarantee of optimality, such algorithms can provide quick solutions, thereby allowing planners to try multiple scenarios. To emphasize maximizing the per-lightpath capacity, while reducing the spectral usage for multi-period planning, a multi-objective combinatorial optimization-based solution was introduced [85]. With the ability to achieve a near-optimal solution to an NP-hard problem in polynomial time, this approach is used as a baseline for conducting network planning studies in this work.

A recent work [86] also provides upgrade strategies for operators to move to L and S-band transmission, thereby providing a network planning tool to undertake RCSA. However, a clear comparison with existing C-band alternatives is missing, which the proposed contribution seeks to address.

4.1.2 Comparison of RCSA alternatives

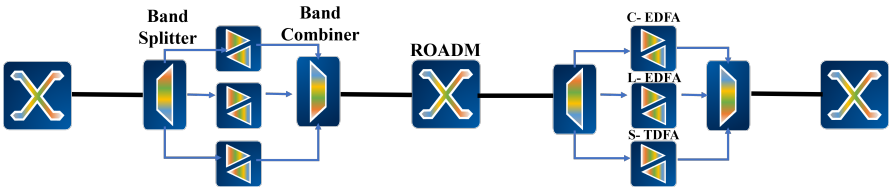


Figure 4.1: Architecture of a multi-band OTN transmission.

Let us consider as baseline the RCSA proposed in Section 4.3 (see Fig. 3.6), which aims at minimizing the number of lightpaths added to the network for each planning period t, while trying to meet the requested aggregated data-rate generated by the same traffic model which caters for uncertain increase. This RCSA considers two approaches: *LP upgrade* (i.e., upgrading the configuration of the BVT if possible, to cope with the required data-rate), and *LP addition* (i.e., adding more BVTs so that the requested data-rate is met) [2].

This RCSA is further extended to achieve a higher capacity without undertaking band upgrade. The proposed solution referred as "C-band", aims at first minimizing the required BVTs and then maximizing the data-rate for each demand $d \in D$ in planning period t, while trying to meet the requested aggregate data-rate. For this purpose, three different approaches were considered: *LP upgrade*, *LP addition*, and *LP rerouting* (i.e., a spectrum re-allocation

[2] ref. Section 3.2

heuristic to reroute spectrally adjacent BVTs such that the released spectrum can be used to upgrade the configuration of the placed BVT) [87].

The use of regeneration has been encouraged to cope with long lightpaths as they split the transparent optical distance of the lightpath into several transparent segments[3]. The compromise of using regenerators with respect to the throughput, the number of lightpaths, and other metrics will be analyzed in detail here.

In order to improve the spectrum utilization per link, the C-band RCSA module is extended to include regeneration capabilities (referred to as "C-band (Regen)"). The regeneration locations are pre-calculated for every source-destination demand pair and every configuration on all k-shortest paths, using the minimum receiver OSNR threshold as a configuration selection metric. C-band (Regen) considers the same approaches as its non-regenerator-based counterpart, with the addition of a secondary objective in LP addition, that is, minimizing the number of regenerator locations in the network. Although adding regenerators can lead to higher throughput and reduce under-provisioning as compared to the C-band approach, there is a need to provision two additional BVTs for every regenerator location for each demand.

Band Division Multiplexing (BDM) exploits the use of neighboring bands (C+L-bands and S+C+L-bands in this work), which are a more cost-effective option than space division multiplexing (i.e., lighting dark fibers) [66]. Existing BVTs can be tuned to any of these bands. However, although commercially available wideband EDFAs can be used for both C- and L- bands, S-band requires TDFAs. Moreover, QoT parameters to select configurations for each BVT vary across bands due to higher attenuation on the S- and L- Bands [66].

Scenario	**Attenuation (dB/km)**			**Noise Fig (dB)**			**Frequency Slots (12.5GHz)**		
	C	**L**	**S**	**C**	**L**	**S**	**C**	**L**	**S**
C-band	0.22	-	-	5.0	-	-	400	-	-
C+L-band	0.22	0.24	-	6.0	6.0	-	400	400	-
S+C+L-band	0.22	0.24	0.25	6.0	6.0	7.0	400	400	400

Table 4.1: Transmission parameters for multi-band scenario used in this work.

The C-band solutions consist of a 5 THz frequency bandwidth available in the C-band range (i.e. 191-196 THz). This 5 THz bandwidth is split into 400 frequency slots of 12.5 GHz each. Depending on the configuration, each BVT is allocated the required frequency slots, while maintaining the wavelength continuity constraint across the links traversed by its lightpath. To simulate the availability of additional bands, the C+L-band range consists of 800 frequency slots (i.e., 186-196 THz), and the S+C+L-band range consists of 1200 frequency slots (i.e., 186-200 THz). For fairness, a lower QoT on the S-band and the L-band is assumed, due to the higher amplifier noise figure and attenuation. The basic transmission parameters used

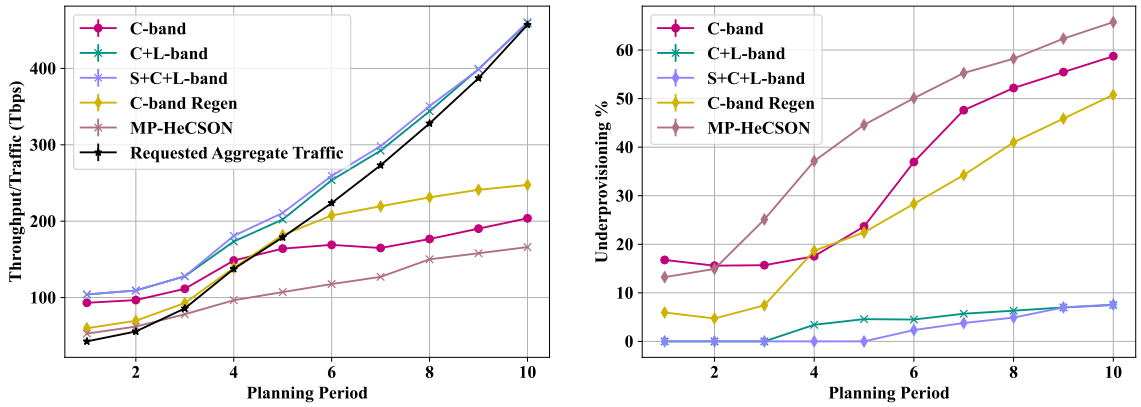
[3]ref. Section 3.3

are shown in Table 4.1. For the band filling strategy in case of multiple bands, the C-band is always filled first, followed by the L-band, and finally the S-band, if available.

4.1.3 Simulation Results and Discussions

The proposed RCSAs have been compared for different networks: Nobel-Germany, Nobel-EU, Abilene-US, and Sweden [30, 12]. These networks differ not only on the average node degree, and the number of demands $\mathcal{D}$, but also on the link lengths, which will have an impact on the configurations that can potentially be used (the longer path, the less suitable configurations). Each link $l \in \mathcal{E}$ consists of a single fiber pair and the transponders are equipped to handle 26 different configurations for the C-band (see Fig. 2.2).

Let us first compare the network throughput for each alternative (referred to as MP-HeCSON, C-band, C-band regen, C+L-band, S+C+L-band) with respect to the requested data-rate, as shown in Fig. 4.2a. It can be observed that only multi-band solutions are able to meet the expected throughput throughout the planning period. Also, the use of regenerators postpones the saturation of the C-band by 3 years. Furthermore, the multi-band solution shows similar throughput and data-rate distributions, pointing out that an S-band may not be needed for this planning.



(a) (b)

Figure 4.2: Simulation results and comparison. (a) Network throughput and (b) underprovisioning percentage for Nobel-EU over 10 planning periods.

To fairly measure the number of demands which could not meet the requested traffic, we introduce a metric called as “underprovisioning". For a given planning period t, underprovisioning is defined as the ratio of the sum of all demand data rates which could not be provisioned in the network to the overall requested aggregate traffic. In this work and subsequent sections, underprovisioning UP_t is represented as a percentage, and can be calculated as shown in Eq. 4.2:

$$UP_t = \frac{\sum_{\forall \widetilde{d} \in \widetilde{D}_t} (DR_{\widetilde{d}} - \sum_{\forall l \in LP_{\widetilde{d}}} DR_l)}{\sum_{\forall d \in \mathcal{D}_t} DR_d} \cdot 100\ \% \tag{4.2}$$

where $\widetilde{D}_t$ is the set of demands in planning period t which could not be satisfied by the placement of in-operation lightpaths $l \in LP_{\widetilde{d}}$. $\mathcal{D}_t$ is the set of all demands in the planning period t.

Fig. 4.2b depicts the underprovisioning as a percentage for the different solutions. Although all the solutions showed to offer more throughput than the requested one for the first two years, Fig. 4.2b shows that none of the C-band solutions cope with the 100% of the demands for any year. For example, in year 2, "MP-HeCSON" and "C-band" could not cope with 15% of the data-rate, but the use of regenerators reduced the under-provisioning to 5%. In these cases, this under-provisioning was compensated by the overprovisioning of other demands, hence, increasing the total throughput above the requested one as shown in Fig. 4.2a. The C+L-band and S+C+L-band solutions defer under-provisioning 4 and 7 years respectively, always keeping it lower than 8%.

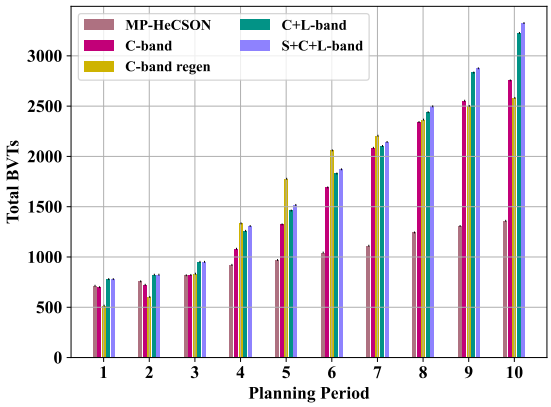


Figure 4.3: Total BVTs deployed in the Nobel-EU network for each of the scenarios over 10 planning periods.

Another important criterion for operators is the number of required BVTs as it will have a direct impact on the expected capital and operational expenditures (including power consumption). This comparison is depicted in Fig. 4.3, which shows that the higher the throughput, the more lightpaths are set and hence, the more BVTs are deployed. The number of lightpaths keeps increasing for all the solutions, especially when using regenerators. This fact may seem inconsistent for the C-band solutions, which are not able to provide the expected throughput (Fig. 4.2a) but they can still set lightpaths through non-congested links. In year 8, using regenerators in the C-band needs similar number of BVTs as the BDM solutions (close to 2500 BVTs), but offering less throughput (200 Tbps instead of 300 Tbps) and more

under-provisioning (33% instead of 6%). Focusing on the C-band, regeneration increases the throughput and reduces the under-provisioning but requires more BVTs.

4.1.4 Summary and Conclusions

This work presented and compared different RCSA solutions in a multi-period planning scenario in a single fiber-pair use-case. The results of network planning solutions on the Nobel-EU network topology show the need of using multi-band solutions in order to meet the requested aggregate data-rate. When operators want to restrict themselves to the C-band, the use of regenerators can increase the network throughput with lower under-provisioning as compared to its non-regenerator counterpart.

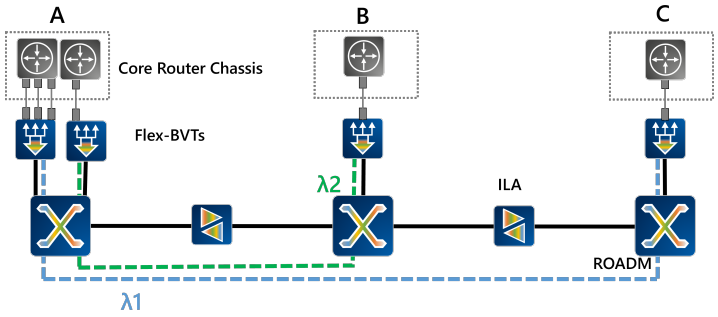
4.2 Network Upgrades for Multi-band Planning

RCSA algorithms for long-haul OTNs not only serve as a tool to plan new services into the networks, but also to conduct forward-looking strategic studies, thereby providing stakeholders with a cost and effort estimate. One such study is using combinatorial optimization-based RCSA algorithms to find the performance of different terminal equipment upgrades.

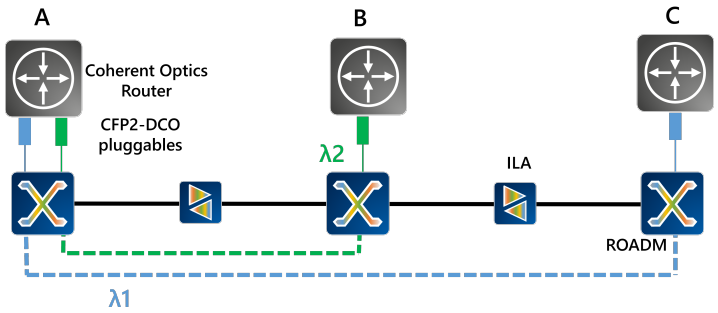
As introduced in Section 1.1.5, OTN operators have two choices for terminal equipment deployment in their networks, namely BVTs and CFP2-DCO. BVTs aggregate Ethernet client signals onto a single coloured interface, before launching it as a lightpath into the OLS. CFP2-DCO on the other hand, are IPoWDM-based pluggables that can provide coloured interfaces directly to the core IP routers. This allows the creation of lightpaths from the routers directly to OLS, thereby eliminating the need of using terminal aggregation devices. Fig. 4.4 shows the difference in architecture between a BVT and CFP2-DCO based architecture.

As seen in Fig. 4.4a, the advantage of the Flex-BVT deployment strategy is that the operator can re-use most of the Layer 3 core routers, since BVTs can perform O-E-O as an intermediary between the core routers and the OLS. Further, the OTN operator is no longer restricted to choosing coherent optics-based routers, leading to wider flexibility in the choice of routers.

In Fig. 4.4b, a multi-haul CFP2-DCO-based transparent IPoWDM architecture achieves data rates up to 400 Gbps by eliminating the optical terminal and using coherent optics-enabled core routers which have CFP2-DCO cages. Each of these CFP2-DCO pluggables can be tuned to a specific central channel frequency on the C-band and launched directly into the OLS. The advantage of such a solution is that the network operator does not need to invest in additional optical terminals, since the O-E-O conversion happens on board the pluggable, while the core



(a) Flex-BVT-based DWDM and IP layer network architecture



(b) CFP2-DCO-based transparent IPoWDM architecture

Figure 4.4: Flex-BVT versus CFP2-DCO deployment for a simple three nodes and two links, deploying two lightpaths $lambda_1$ and $lambda_2$.

router takes care of the electrical grooming of client interfaces. The disadvantage, however, is that the network operator must deploy expensive core routers which can provide enough power and backplane switching capacity to support such pluggables.

Traditionally, an IPoWDM deployment consists of network deployments that use Quad Small Form-factor Pluggable Direct Detection (QSFP-DD)-based pluggables. These pluggables have the capability of achieving up to 400 Gbps and can be integrated into QSFP-DD cages on core IP routers. However, due to the smaller footprint and reach limitations of these pluggables, they are restricted to metro and data-center applications. A long-haul network with QSFP-DD pluggables can be realized by using an opaque deployment, which has several disadvantages, both from a techno-economic and operations perspective [88]. On the other hand, multi-haul CFP2-DCO pluggables are a more suitable option to realize transparent IPoWDM networks, since they provide similar reach as compared to a BVT, but are less expensive. Therefore, it is pertinent that the suitable IPoWDM pluggable is chosen, based on the type of networks and requirements of the operators [89].

While BVT-based and CFP2-DCO-based transparent IPoWDM architectures have their advantages and disadvantages, network operators must evaluate both scenarios and invest accordingly.

In this contribution, a multiband network planning study is conducted to compare the effects of terminal devices on band activation. The first scenario is when only BVTs are used, whereas the second scenario is when only CFP2-DCOs are used. A CAPEX study for the terminal and OLS upgrades is also undertaken in both the cases thereby highlighting the advantages and disadvantages of BVT and CFP2-DCO.

4.2.1 State-of-the-Art Analysis

Since BDM is a CAPEX-intensive activity owing to upgrades at each ILA location in the OLS, the research on increasing optical network capacity has primarily focused on the development of BVTs. An architecture of a BVT that could achieve up to 400 Gbps 16QAM long-haul transmission was proposed [90]. With further advances in DSP, and the standardization of ITU-T flex-grid frequency slots, the interest in conducting network studies exploiting such characteristics also increased.

In recent years, IPoWDM-based solutions, namely QSFP-DD and multi-haul CFP2-DCO have emerged as a competitor to BVT-based solutions [91]. Within IPoWDM, there exist two methods of network design, namely, an opaque network design (using QSFP-DD-based pluggables) and a transparent network design (using multi-haul CFP2-DCO). For long-haul core networks, network studies were conducted based on the number of transceivers, and multi-haul CFP2-DCO emerged to provide a better cost per throughput in such networks [88, 89]. However, these studies ignore the fact that most core IP routers do not have CFP2-DCO cages, which would force network operators to invest in routers that support multi-haul CFP2-DCO-based coherent routing [92].

Multi-layer network planning studies with a flex-grid deployment proved to hold several advantages over fixed-grid deployments [93, 94]. A techno-economic planning study with BVTs up to 400 Gbps 16-QAM was also undertaken, which showed more than 50% savings on equipment cost as compared to fixed-grid deployments [68]. However, most of these studies were aimed at maximizing the C-band capacity, with minimal focus on realistic network traffic increase, and the option of using frequency slots outside of the C-band.

Therefore, this work undertakes a network operator-centric multi-band multi-period planning study to increase the offered traffic in single-mode fiber long-haul OTN deployments. To this end, two terminal upgrade solutions are compared on a network OLS which allows C, L, and S-band transmission. The terminal upgrade solutions are compared not only on their ability to achieve the requested traffic, but also on how early the new transmission bands are activated. Finally, network deployment CAPEX costs are evaluated using a simple cost model to compare the costs of BVT and multi-haul CFP2-DCO deployments.

4.2.2 Routing and Configuration Pre-Selection

Before the periodical planning of lightpaths for the given traffic request begins, two major steps are taken. These are, routing of source-destination pairs, and a pre-selection of candidate lightpath configurations, based on the optical performance of the given BVT/CFP2-DCO in all the bands of operation. Algorithm 1 (Appendix C) gives an overview of this process.

For the routing, $k = 3$ Dijkstra's shortest path non-disjoint routing is calculated for each source-destination pair in the demand matrix $\mathcal{D}$. It is observed that in long-haul OTNs the path diversity increases drastically for higher values of k, thereby leading to higher delays and lower data rates per transparent lightpath. Once the source-destination pairs are assigned their routes, they are ordered according to their first shortest path lengths, with the source-destination pair having the longest route being assigned the highest priority. Then, source-destination pairs with similar path lengths ($\leq$ 100 km) are ordered according to the number of hops, with the pairs having a higher number of hops being assigned a higher priority. This is done to introduce fairness in the simulation so that the pairs using more links are not blocked later by pairs using fewer links but more spectral resources.

Next, using *HeCSON* (ref. Section 3.1), candidate BVT/CFP2-DCO configurations are found for each source-destination pair in the network. For each possible configuration, *HeCSON* simulates a fully-loaded homogeneous channel and finds the worst-case channel GSNR using the modulation format dependent Inter-Sysmbol Stimulated Raman Scattering Gaussian Noise (ISRS-GN) model to calculate the NLIN term [33]. The ISRS-GN model is used here, because of its capabilities to estimate NLIN for central channel frequencies in a multi-band scenario. To ensure that the NLIN in the L and S bands is not underestimated, additional GSNR margins are added to the calculated GSNR, as given in [95]. For each routed demand, the configurations whose worst-case calculated GSNR is lesser than the minimum receiver sensitivity GSNR are removed from the list of possible configurations. After this, the candidate lightpaths which can satisfy the traffic requirement for each of the demands are found.

4.2.3 Selecting Candidate Terminal Device Configurations

Once the paths and configurations of all source-destination pairs and the traffic matrix for the current planning period τ_t (ref. Eq. 2.9) is calculated, for each demand, suitable configurations are assigned for each network terminal card, such that the traffic demand between each source-destination pairs can be satisfied with minimum equipment.

Let us consider a demand $d(i, j) \in \mathcal{D}$ in planning period t, with additional requested traffic $\delta(i, j, t) \in \tau_t$ (ref. Eq. 2.14). It may be the case that some of the additional requested

traffic may already be satisfied by some in-operation lightpaths. Therefore, the traffic to be met between source i and destination j of demand d at a given time t is defined as $\rho(i,j,t) = \tau(i,j,t-1) + \delta(i,j,t) - \sum_{\forall l \in \widetilde{L_d}} DR_l$, where $\tau(i,j,t-1)$ is the total requested traffic up to the previous planning period for demand d, and $\sum_{\forall l \in \widetilde{L_d}} DR_l$ is the sum of the data rates DR_l of all the in-operation lightpaths $\widetilde{L_d}$ belonging to demand d.

Using the above definitions, the selection of candidate lightpaths from the demand's possible configurations $C_d \in C_{\mathcal{D}}$ is defined as,

$$\text{Minimize} \sum_{c \in C_d} \sum_{p \in SP_d} x_{c,p}, \text{Maximize} \sum_{c \in C_d} \sum_{p \in SP_d} \left(x_{c,p} \times DR_c\right), \tag{4.3}$$

$$\text{such that } : \rho(i,j,t) \leq \sum_{c \in C_d} \sum_{p \in KSP_d} x_{c,p} \times DR_c \leq \rho(i,j,t) + \eta, \tag{4.4}$$

$$\sum_{c \in C_d} x_{c,p} \times BW_c \leq \widetilde{BW}_p, \forall p \in SP_d, \tag{4.5}$$

$$\text{integer variables } x_{c,p} \geq 0, \forall c \in C_d, \forall p \in SP_d.$$

In Eq. 4.3, a multi-objective integer linear objective function is defined, which minimizes the number of candidate lightpaths (integer variable $x_{c,p}$) while maximizing the data-rate DR_c of each of the candidate lightpaths associated with configuration $c \in C_d$ along path $p \in SP_d$.

This objective function is constrained by Eq. 4.4, known as the data-rate constraint. η is a constant which limits the overprovisioning of the data-rate, avoiding greedy usage of spectral resources.

In Eq. 4.5, a bandwidth constraint is introduced for each path $p \in SP_d$, such that the total bandwidth of all candidate lightpaths on path $p \in SP_d$ shall not exceed the maximum available contiguous bandwidth on the path p, denoted by $\widetilde{BW}_p$.

Using Eq. 4.3-4.5 allows an optimal selection of candidate lightpaths which can be then placed onto the spectrum. This operation is continued till all the demands in the given planning period are satisfied.

4.2.4 Multi-period RCSA algorithm

In a multi-period planning scenario, the heuristics and ILPs previously discussed in this section are used to place lightpaths into the network, to cope with the per-period increase in data-rate demands between any given source i and destination j in the network.

Algorithm 2 provides an overview of the described RCSA algorithm in a multi-period scenario. As a pre-requisite, the candidate paths of demands $\mathcal{D}$ and their corresponding candidate configurations $C_{\mathcal{D}}$ should already be pre-calculated using the routing and configuration

pre-selection algorithm (ref. Alg. 1). For each planning period $t \in \mathcal{T}$, each demand $d \in \mathcal{D}$ is iterated over to find the source i and destination j belonging to it.

Once the additional requested traffic $\delta(i, j, t)$ is calculated, the in-operation lightpaths $L_d \in \mathcal{L}$ need to be upgraded. The upgrade occurs without using additional spectral resources, and without changing the center channel frequency of each in-operation lightpath. The upgraded in-operation lightpaths for the demand d are denoted as $\widetilde{L_d}$. Further details on the lightpath upgrade heuristic are available in our previous work [4].

Using the data-rate of the upgraded lightpaths for demand d, $\rho(i, j, t)$ is calculated as described in the previous section. If the upgraded in-operation lightpaths are unable to satisfy the additional requested traffic (i.e. $\rho(i, j, t) > 0$), Eqs. 4.3-4.5 are invoked to find the optimal candidate lightpaths for the demand. The candidate lightpaths $L_{d,t}$ are added to the spectral slots available on the links in a first-fit fashion. The newly placed lightpaths are denoted as $L"_{d,t}$. In case some lightpaths cannot be placed, the re-routing algorithm is invoked [85]. Since the in-operation lightpaths change their bandwidths or central channel frequencies, the GSNR for all the in-operation lightpaths sharing the same links as $L"_{d,t}$ needs to be re-calculated. The remaining traffic for demand d, given by $\rho(i, j, t)$ is updated by subtracting the data-rate of the placed lightpath $L"_{d,t}$. Finally, in case there is traffic remaining that could not be assigned to a lighpath, the underprovisioning percentage $\mathcal{U}$ is increased using Eq. 4.2.

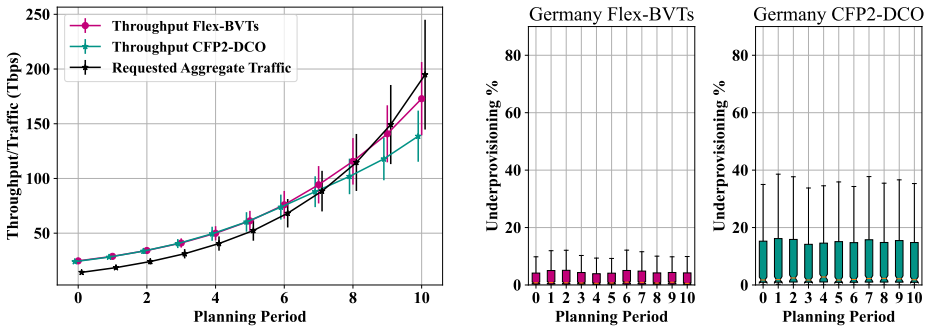
4.2.5 Results and Discussions

The results of the two deployment simulations are compared in terms of aggregate network throughput, number of lightpaths deployed, underprovisioning, and lightpaths per band of operation. Also, the two terminals are compared in terms of cost-per-bit using a "cookie-cutter" techno-economic analysis approach. The network studies are conducted on three different European OTNs, namely, Nobel-Germany, Spain, and Sweden.

Aggregate throughput is defined as the sum of the data rates of operational lightpaths in the network at a given planning period. As mentioned earlier, a lightpath is an end-to-end optical signal between a pair of transceivers (BVTs or CFP2-DCOs) at the source and destination, each having its own configuration. The aggregate throughput is compared with the requested aggregated traffic, which is defined as the sum of the demands between each source-destination pair in a given planning period.

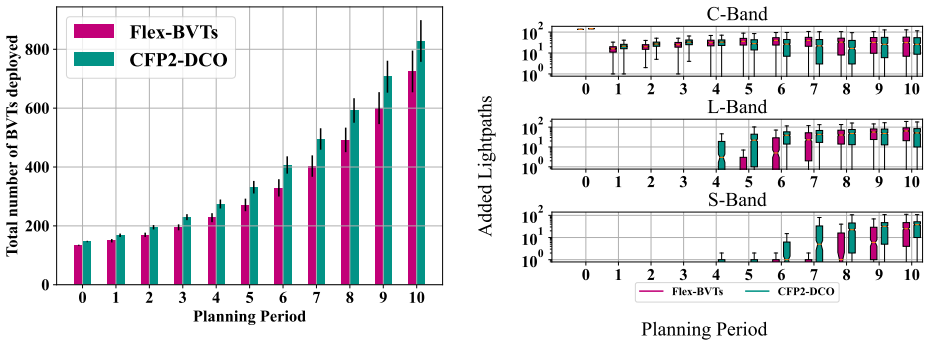
The number of lightpaths is the cumulative count of in-operation lightpaths deployed in each planning period, and finally, the number of lightpaths per band shows how many lightpaths are added in the C, L, or S-band in each planning period. Using these four metrics, and the same RCSA algorithm, a multi-period planning study comparing the performance of

BVT and CFP2-DCO on the three networks under study is carried out. The results for each of the networks are presented and discussed in detail.



(a) Aggregate throughput (b) Underprovisioning

Figure 4.5: Throughput and Underprovisioning results for Nobel-Germany network.



(a) Total deployed lightpaths (b) # lightpaths per band

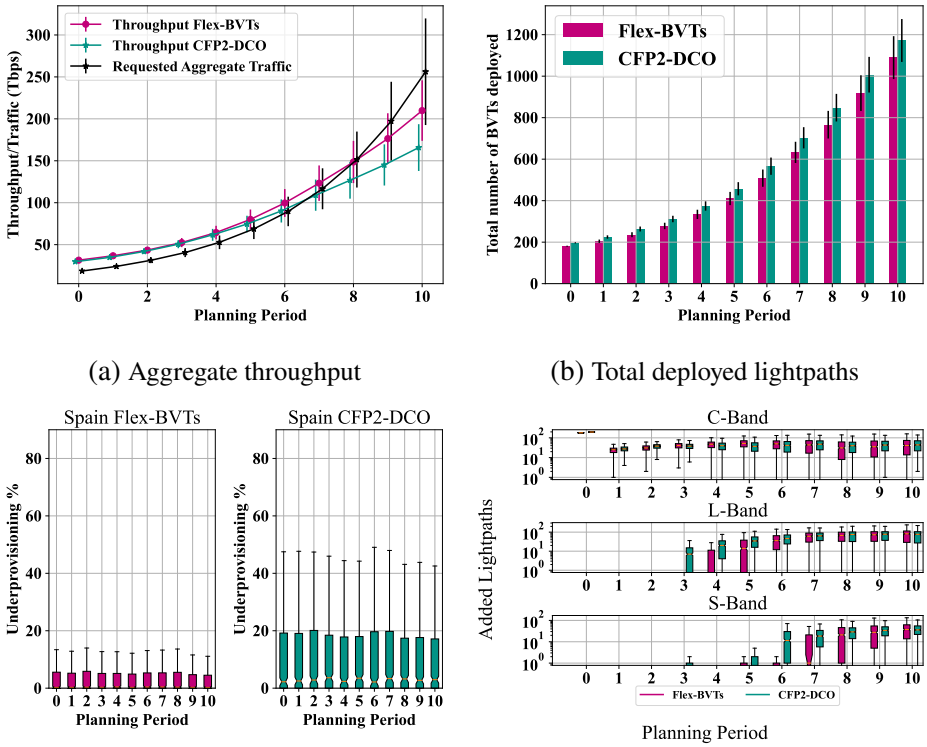
Figure 4.6: Number of deployed lightpaths and the number of lightpaths in each of the bands for the Nobel-Germany network.

From the throughput plots for the Nobel-Germany network in Fig. 4.5, it is observed that the BVT approach meets the requested aggregate traffic one year more as compared to the CFP2-DCO approach, with similar mean underprovisioning. However, CFP2-DCO deployment shows a higher third and fourth quartile in each planning year. This means that it is susceptible to having more lightpaths blocked due to high neighbouring channel interference, or due to band saturation. Compared to the CFP2-DCO approach, the BVT approach places at least 15% lesser lightpaths in the network, leading to a lesser number of devices.

As seen in Fig. 4.6b, S-band is activated two years later in the BVT scenario, as compared to the CFP2-DCO scenario. The reason why the CFP2-DCO approach activates the L and S band in the fourth planning period is that the increase in traffic leads to high usage of spectral resources on a few links which carry many lightpaths (e.g. Frankfurt-Leipzig, Hannover-Frankfurt, and Hannover-Leipzig), leading to L-band saturation in one planning period. This

phenomenon is not observed for BVTs due to the usage of configurations with higher spectral efficiency.

From the results of the network study for Nobel-Germany, it can be inferred that the BVT approach provides a similar or higher throughput than CFP2-DCO-based transparent IPoWDM deployment. If only network metrics are concerned, both solutions are comparable to each other.



(a) Aggregate throughput

(b) Total deployed lightpaths

(c) Underprovisioning

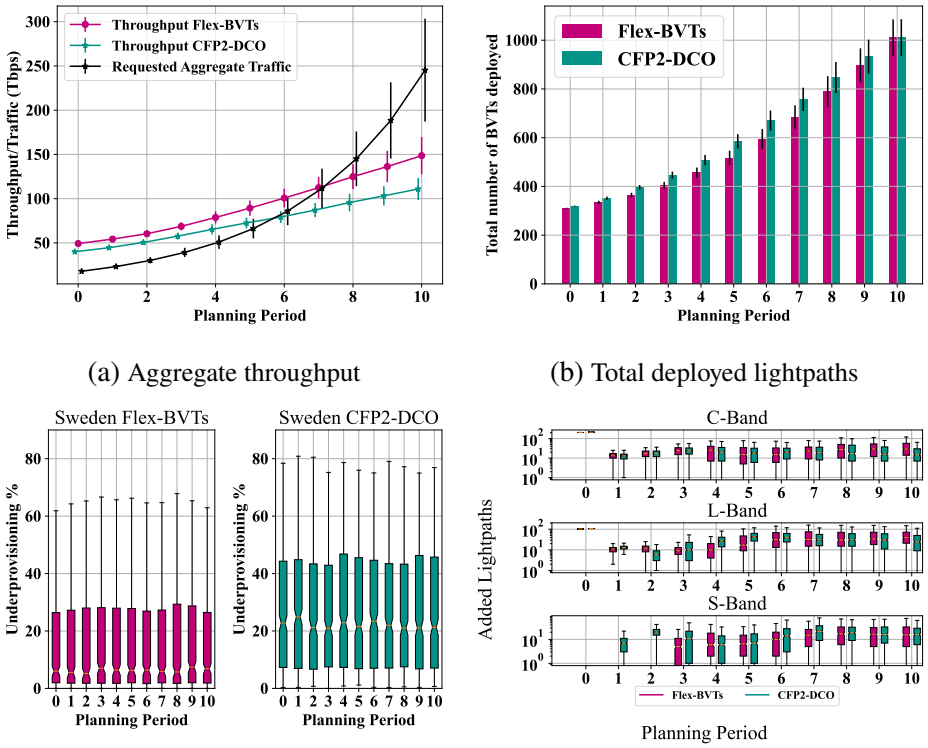
(d) # Lightpaths per band

Figure 4.7: Ten periods planning results for the Spanish network.

The Spanish network results, shown in Fig. 4.7 also highlight the advantages of the BVT approach with higher throughput and lower underprovisioning, as compared to the CFP2-DCO approach. In Fig. 4.7a it is seen that the BVT-based architecture allows for an additional planning year above the requested traffic. This is due to the star-like structure of the network, the links connecting the node Madrid get saturated earlier when using the CFP2-DCO approach. Due to this blocking, the CFP2-DCO approach shows a higher underprovisioning in Fig. 4.7c as compared to the BVT approach. Finally, as shown in Fig. 4.7d, the BVTs used instead of the CFP2-DCO approach, the L-band and S-band activation can be postponed by 1 and 2 planning periods respectively.

Despite the change in topology, higher traffic, and longer average lightpath lengths (see Table 2.1), the BVT approach can provide 20% higher throughput, while placing approximately

10% lesser lightpaths on an average. Moreover the delay of activating the L-band and S-band is always advantageous for OTN operators, since upgrading to a multi-band OLS is a cost-intensive exercise.



(a) Aggregate throughput

(b) Total deployed lightpaths

(c) Underprovisioning

(d) # lightpaths per band

Figure 4.8: Ten periods planning results for the Swedish network.

For the Swedish network, Fig. 4.8 shows a 15% higher underprovisioning on average for the CFP2-DCO approach, as compared to the BVT approach. In terms of throughput, both approaches can only meet the requested aggregate traffic up to the sixth planning period. Due to the longer average lightpath length in the networks and the point-to-point link structure of the network (Fig. 2.3), many demands cannot be satisfied and several links need to be upgraded to a multi-fiber setup, which is out of the scope of our network studies.

However, if the BVT approach is used, S-band deployment in the Swedish network can be postponed by three planning periods, as compared to the CFP2-DCO approach.

As seen from the results, a BVT-based deployment outperforms the CFP2-DCO-based deployment for all the three core networks under study, highlighting its benefits of higher flexibility in configuration selection, higher spectral efficiency, and lower number of deployed devices in long-haul transparent optical networks.

In business parlance, a ‘cookie-cutter’ approach is sometimes used to generalize steps, streamline processes and maintain uniformity. In a techno-economic analysis for networks,

a similar approach is employed to arrive at estimates quickly and to give users an overview of the costs which need to be added to the network. Although not a substitute for careful and deliberate planning applied in the industry before creating commercial offerings, a quick estimate also helps in fastening the decision process by eliminating solutions with high costs to benefits ratio.

In this work, since two different terminal deployments, and adding two additional bands are compared, the costs of adding extra equipment into the network need to be studied. The component cost values are provided in Hernandez *et. al.* [96] and it is assumed that the presently deployed network consists of a C-band OLS, along with pre-existing core IP routers. It is important to note that the costs are normalized according to the cost of a 10 Gbps transceiver in the year 2020.

The OLS upgrade can be planned in two steps. The first step is the replacement of pure C-band EDFAs with ILA modules consisting of C-band and L-band EDFAs along with the placement of passive band splitters and combiners at each location. The second step is the installation of ILA modules consisting of S-band TDFAs. This two-step approach enables operators to install and operate S-band-based ILAs, only when a need for the same arises [97, 98]. Here, a cost sensitivity analysis is not conducted, since this deviates from the cookie-cutter philosophy and requires taking into consideration component customization. Therefore, for the CAPEX calculations, mean values from the results shown in Figs. 4.5-4.8.

As seen in Table 4.2, CAPEX costs related to routers will increase, as the overall throughput in the network increases. It is clear from the throughput results of the three networks, that the BVT-based solution consistently places more throughput than the CFP2-DCO-based solution, thereby leading to a higher overall CAPEX. However, it must be noted that higher throughput would also lead to the ability of the network operator to offer services to more customers. Therefore, the yearly and cumulative CAPEX of each solution is normalized to the carried traffic.

Component	**CAPEX (normalized C.U.)**
BVT (100-600G)	17
CFP2-DCO Transceiver Card (100-400G)	12
Non-coherent IP Router	12 per 400G
Coherent optics IP Router (with CFP2-DCO cages)	20 per 400G
C, L-band EDFA	2.8
Band splitter/combiner pair	0.2
S-band TDFA	5.6

Table 4.2: Cookie-Cutter CAPEX values consolidated from [96, 99] for terminal and OLS equipment upgrade.

For this, two cost-per-bit metrics are defined. The first metric, whose results are shown in the upper halves in Fig. 4.9 and Fig. 4.10 is the yearly CAPEX per yearly added traffic.

This is simply the cost of new equipment in Cost Units (C.U.) normalized to the additional throughput added to the network in Tbps. The second metric, seen in the bottom halves of the same figures shows the cumulative CAPEX normalized to the total offered throughput of the network. The cumulative CAPEX in any given planning period is defined as the CAPEX costs accumulated from the first planning period to the current one. As the planning periods progress, it is possible that many transceivers can be upgraded to higher data rates to meet the requested traffic, thereby avoiding the placement of new terminal equipment in the network.

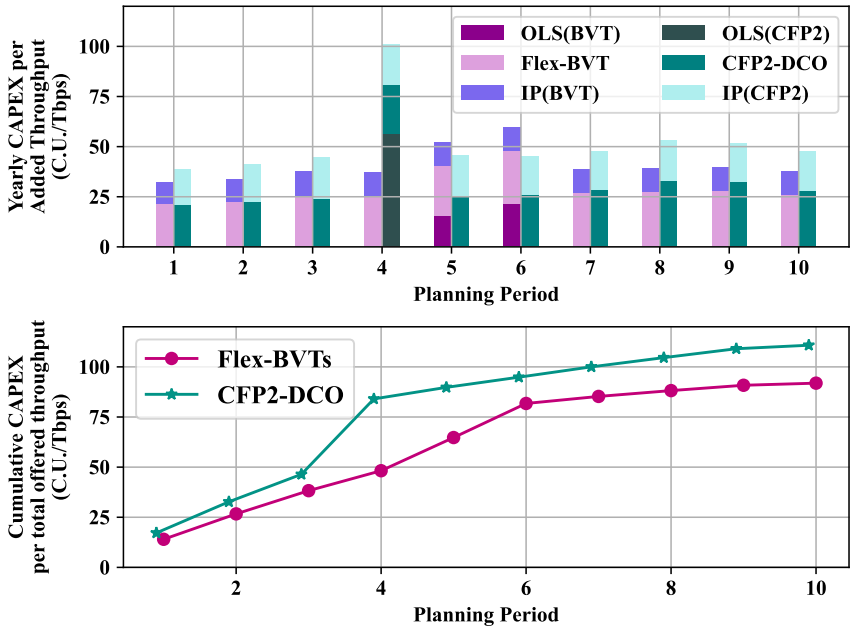


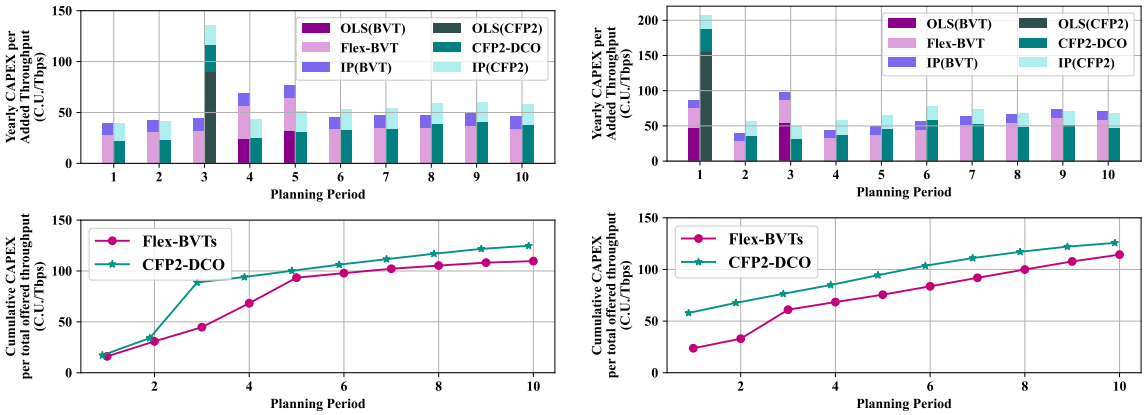
Figure 4.9: Per-period and cumulative CAPEX normalized to per-period and total offered throughput respectively, for Nobel-Germany network.

As seen from Fig. 4.9, the per-component CAPEX normalized by the additional throughput added to the network shows that the BVT solution has a lower cost per bit as compared to the transparent IPoWDM-based solution. Although the costs of the terminal equipment are higher in the case of BVT, they are offset by the usage of less expensive routers. Moreover, the absolute cost of OLS upgrade for both BVT and transparent IPoWDM scenarios is the same. However, since transparent IPoWDM carries lesser traffic, the cost per bit value turns out to be higher than the BVT case.

Specifically, at the end of the planning periods, the usage of the BVT solution provides an overall 12% reduction in the cost-per-bit, as compared to the transparent IPoWDM solution. More importantly, the cumulative CAPEX per bit is consistently lower in the case of the BVT solution, showing that if the offered traffic by the network operators is used to its maximum, BVTs indeed show promise in terms of higher flexibility, and lower cost-per-bit. As seen in

Fig. 4.10, similar savings are seen in other networks, with the Spanish network showing an 11% reduction in the cost-per-bit and the Swedish network portraying a 6% reduction in the cost-per-bit.

This exercise helps not only to compare the differences in the cost per bit but also the benefits offered by the OLS and terminal upgrades. Higher IP routing costs are assumed due to the usage of coherent IP routers with CFP2-DCO cages as compared to grey-optics IP routers which can be either re-used or acquired at lesser costs as compared to their coherent counterparts [100].

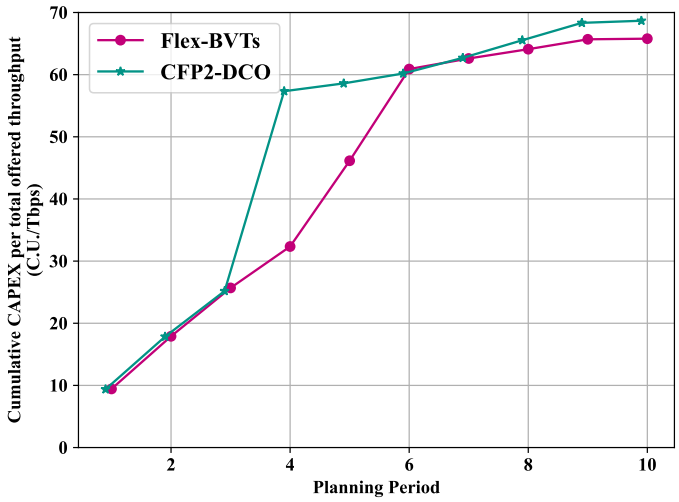


(a) Spain (b) Sweden

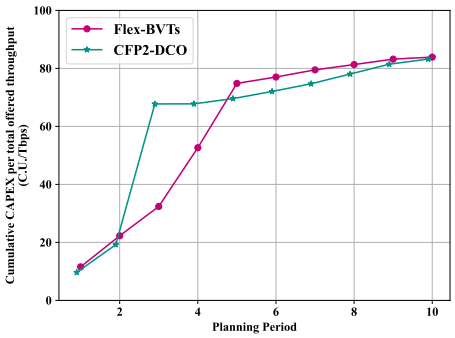
Figure 4.10: Per-period and cumulative CAPEX normalized to per-period and total offered throughput respectively, for Spanish and Swedish networks.

For fairness in cost comparison, a CAPEX study with only the OLS and Terminal costs is also conducted. This is done to provide more insights into the cost components purely from a WDM perspective. For the three networks under study, a cost-per-bit analysis is also undertaken, using the costs provided in Table 4.2, but only considering the cost of OLS upgrades and additional terminals. The metric for comparison is yearly CAPEX per total offered throughput in the network, which is measured in C.U. per Tbps.

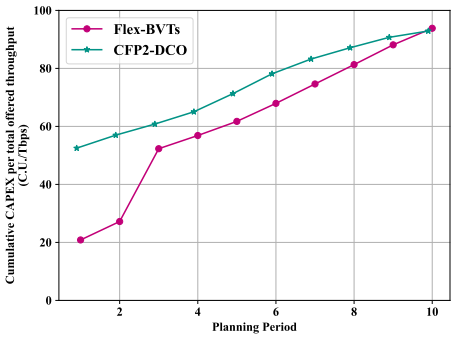
From Fig. 4.11, it is seen that in most of the planning periods in all three networks, the BVT approach is either similar to or less expensive than the CFP2-DCO approach. The exceptions are planning periods 5-9 in the Spanish network and planning period 10 in the Swedish network. The reason for a lower CAPEX-per-bit for the CFP2-DCO approach is that from the fifth planning period, the L-band starts saturating, leading to a lesser number of CFP2-DCO transceivers being added to the network, thereby also having an overall lower amount of throughput. However, towards the end of the planning periods, both the curves begin to converge, showing little to no economic advantage of investing in CFP2-DCO technologies for a long-haul transparent optical network deployment.



(a) Nobel-Germany



(b) Spain



(c) Sweden

Figure 4.11: Cumulative CAPEX per offered throughput for the three networks under study, ignoring IP routing costs.

4.2.6 Summary and Conclusions

As research communities in Europe move towards standardization of 6G network services and deployment, an exponential increase in demands for high-bandwidth services to be carried over long distances is envisioned. Moreover, as 5G networks are already partially deployed in production environments, long-haul network operators must take timely decisions to upgrade core networks. These network upgrades can either be related to the choice of optical network terminal equipment or the choice of the type of OLS upgrades needed in the network.

Using the traffic model provided in Section 2.3 and the RCSA algorithm introduced, a multi-period planning study on the German, Spanish, and Swedish networks is conducted. This study compares differences between multi-band deployments based on highly flexible bandwidth variable (BVT) transceivers, versus transparent IPoWDM deployments that use CFP2-DCOs. Results show that the BVT approach places up to 15% lower lightpaths into the network while carrying up to 20% higher traffic. Using a 'cookie-cutter' CAPEX calculation

approach, it is also shown that the cost-per-bit of deploying bandwidth variable transceivers is up to 12% lower, as compared to the CFP2-DCO approach under study.

Additionally, a cost-per-bit study comparing the two approaches only from the OLS and terminal costs perspective shows that the BVT approach has either similar or lower cost per bit in 80% of the planning periods for the three networks in the study.

4.3 Migration Strategies for Multi-band OTNs

Previous sections have extensively discussed several upgrades to the OTN, whether it be an upgrade of the terminal equipment to increase spectral efficiency and achieve higher data-rates, or an upgrade of the OLS to accommodate traffic via multiple fibers or multiple bands.

These upgrade (or migration) decisions need to be optimized firstly, in order to meet the traffic increase, and secondly, in order to minimize the upgrade costs for the OTN operators. This contribution evaluates migration strategies to either migrate to a new band or to a new fiber pair on the links of an OTN. For the band upgrades, the C, L, and S-bands are chosen and for each link, up to four fiber pairs can be added to the OTN simulation. The two upgrade strategies implemented are an ILP solution and a heuristic-driven RL solution. These two solutions are compared extensively with each other, highlighting their advantages and disadvantages.

The upgrade strategy obtained from both these solutions is then associated with a cost model and a simple CAPEX-oriented techno-economic analysis is performed, thereby providing OTN operators an initial estimate of the cost.

4.3.1 State-of-the-Art Analysis

Several studies have investigated multi-band provisioning and upgrade strategies for OTNs. Specifically, ML-based models have also been explored in order to improve heuristics and reduce upgrade times.

A RCSA algorithm along with a regenerator placement solution was presented by Sadeghi *et. al.*, which can place lightpaths on multiple bands and fibers [84]. However, the suggestion of the usage of regenerators for S-band lightpaths makes the solution unrealistic for future deployments. A multi-band provisioning scheme was presented by Sambo *et. al.*, which uses k-shortest paths and GSNR-based path verification and downgrades BVT configurations in case GSNR is violated [90]. However, the solution does not take into account multiple fiber pairs in the network due to which significant demand blocking is observed.

Morales *et. al.* present a RL-based solution for the multi-band network provisioning scenario [101]. They also evaluate different agents with the stable baselines framework and provide insights into the applicability of RL into the network. Although significantly advanced for its time, the solution does not optimize or change the reward function and is applicable only to a single fiber pair scenario, rendering it limited in terms of real-world applicability. Finally, Sheikh *et. al.* implement a Deep Q-learning-based solution for multi-band OTN and test it on two differently sized network simulations. However, the agent does not cater to multi-fiber upgrades and the RL cannot outperform heuristics in terms of blocking probability.

4.3.2 ILP-based Upgrade Strategy

This work aims to obtain the best migration strategy across the OTN planning horizon. Each year, the network selected will request an amount of traffic per demand (generated using Eq. 2.7). The demands are ordered from the highest traffic requested to the lowest, allowing for a priority on the requested traffic demands. The objective of the ILP is then to find the best combination of band and fibers to allocate the candidate lightpaths for all the demands in the given planning period. Two segments will be presented to explain the ILP solution. The first refers to the parameters and variables of the problem. Parameters are given inputs that allow the definition of the constraints and do not change. The variables are the outputs of the ILP that define which upgrades are needed for every link in a given planning period. The second segment defines the objectives and the constraints of the ILP, in order to achieve converged optimization.

4.3.2.1 Parameters and Variables

There are six parameters/inputs to the ILP. First, the parameter τ^d corresponds to the total requested data rate for a given demand in a given planning period. Then, the parameters $Slots^d_{p,b}$, $Lightpaths^d_{p,b}$ and $DR^d_{p,b}$ indicate how many slots, lightpaths and data rate are needed by a "chunk" of the demand in a path and band. A chunk is defined as a set of homogeneous lightpaths which can carry a part of the requested demand in a given planning period. A demand may consist of more than one chunk. Each chunk has the same number of lightpaths and can occupy a maximum of 173 12.5 GHz slots in the spectrum. The maximum chunk size is treated as a design choice, as it corresponds to half the spectrum slots in C-band. Therefore, all chunks of the demand have the same number of slots, lightpaths, and total data rates. The set of all chunks across all demands is defined as C. The parameter $Stock^l_{b,f}$ contains the total number of empty slots in link l at band b and fiber f in a given planning period. Further, $a^l_{b,f}$ is the variable to be updated with $A^l_{b,f}$ after every planning period. It stores the already active band and fibers in a link. Before computing the ILP solution for any

demand at the first year, the parameter is active in C-band at the first fiber pair of each link. The notations used across throughout the ILP definition are mentioned in Table 4.3. Further, six input variables to the ILP, which are shown in Table 4.4.

Notations	Definition
$\mathcal{D}$	set of demands
C	set of chunks
P	set of paths
$\mathcal{E}$	set of links
B	set of bands
F	set of available fibers
d	index of a demand
c	index of a chunk
p	index of a path
l	index of a link
b	index of a band
f	index of a fiber

Table 4.3: Notations used in the upgrade ILP.

There are 2 output variables in the ILP problem. The first variable is $U^{d}_{c,p,b,f}$ is a binary decision variable which denotes if chunk c, on path p, deployed on band b of fiber f of demand d is activated or not. Lightpaths belonging to the same chunk are deployed in the same path, band, and fiber. The optimal case is to have only one group, but if the amount of slots from all lightpaths of demand is higher than half the total slots in C-band, the logic considers the splitting of the group into 2 or more according to which value guarantees all groups have fewer slots than half the spectrum of a C-band. The second variable is $A^{l}_{b,f}$, which indicates all the bands and fibers activated for a link in order to provision the traffic of the demand. At each ILP run, $A^{l}_{b,f}$ can only remain with the same values or activate more bands and fibers, since it will have a constraint related to $a^{l}_{b,f}$.

Variables	Type	Definition	Input/Output
τ^{d}	Integer	Traffic requested in Gbps by demand d	Input
$Slots^{d}_{p,b}$	Integer	Slots needed by chunk c of demand d in path p and band b	Input
$LP^{d}_{p,b}$	Integer	Lightpaths needed by chunk c of demand d in path p and band b	Input
$DR^{d}_{p,b}$	Integer	data-rate needed by chunk c of demand d in path p and band b	Input
$Stock^{l}_{b,f}$	Integer	Available slots for link l in band b and fiber f	Input
$X^{d}_{p,l}$	Binary	1 if path p from demand d includes link l, 0 otherwise	Input
$a^{l}_{b,f}$	Binary	1 if band b in fiber f is active for link l, 0 otherwise	Input
$U^{d}_{c,p,b,f}$	Binary	1 if demand d uses chunk c at path p, band b and fiber f, 0 otherwise	Output
$A^{l}_{b,f}$	Binary	1 if band b at fiber f is active for link l, 0 otherwise	Output

Table 4.4: Input and output variables used in the upgrade ILP.

4.3.2.2 Objectives and Constraints

There are five objectives, which will be solved in priority order. The first one (Eq. 4.6) is the minimization of the bands upgraded at each link. Every year, the goal is to upgrade the least number of links possible. Since $A_l^{b,f}$ is 1 if a band in the given fiber is active for a certain link, the objective is defined as the minimization of the sum of $A_l^{b.f}$ for all bands(C, L, S), all fibers and all links. The second objective (Eq. 4.7) is to use as minimum an amount of chunks as possible. Since the selection of chunks goes from lowest to highest index, the weight is considered as $c + 1$. The third objective (Eq. 4.8) aims to minimize the number of fibers. Since the index of the first fiber is 0, the weight defined is $f + 1$. The fourth objective (Eq. 4.9) is aligned with the costs of each band, considering C-band with no cost weight, L-band with a cost weight of 1.2, and S-band with a cost weight of 1.5. The weights are set to be proportional to the ratio of the costs of the components needed for each band, with L-band components assumed to be 20% more expensive than C-band and S-band assumed to be 50% more expensive than C-band. Finally, when the minimum of fibers and bands have been selected, the path for the demand is chosen between the 3-shortest paths (Eq. 4.10). The weight $p + 1$ is defined such that the shortest path ($p = 0$) has a weight of 1.

$$\text{Minimize} \sum_{l\in L, b\in B, f\in F} A_{b,f}^{l}, \tag{4.6}$$

$$\text{then minimize} \sum_{d\in D, c\in C, p\in P, b\in B, f\in F} U_{c,p,b,f}^{d} * (c+1), \tag{4.7}$$

$$\text{then minimize} \sum_{d\in D, c\in C, p\in P, b\in B, f\in F} U_{c,p,b,f}^{d} * (f+1), \tag{4.8}$$

$$\text{then minimize} \sum_{d\in D, c\in C, p\in P, f\in F} U_{c,p,0,f}^{d} + 1.2 \times U_{c,p,1,f}^{d} + 1.5 \times U_{c,p,2,f}^{d}, \tag{4.9}$$

$$\text{finally minimize} \sum_{d\in D, c\in C, p\in P, b\in B, f\in F} U_{c,p,b,f}^{d} * (p+1) \tag{4.10}$$

The ILP problem is subject to different constraints. The first constraint defines that at least one band is active in a link. Since the order of activation is C-band, then L-band and finally S-band, the default activation will be the C-band from the first fiber pair in the link. Now, after every ILP run, the bands and fiber activated from a link should remain active. Hence, $a_{b,f}^{l}$ is used as a memory. In Eq. 4.12, all bands and fibers from a link are checked to see if they must also be active in variable $A_{b,f}^{l}$.

$$\forall l \in \mathcal{L} : \sum_{b\in B, f\in F} A_{b,f}^{l} >= 1 \tag{4.11}$$

$$\forall l \in \mathcal{L}, b \in B, f \in F : A^l_{b,f} >= a^l_{b,f} \tag{4.12}$$

The following constraints are defined to maintain the order of activation of bands and fibers. The indexes for the bands are 0 (C-band), 1(L-band) and 2(S-band). The indexes for the fiber pairs traverse from 0 through 3. Eq. 4.13 indicates that the activation of a higher band can only be possible if the band with a lower index in the same fiber is already active. Eq. 4.14 defines that the activation of a fiber pair with a higher index is only possible if the previous fiber pair is already active. Finally, Eq. 4.15 describes that the activation of a new fiber pair (in C-band by default) can only be possible if the previous fiber pair has all the bands active. In this case, to assure all bands active it is only needed to see if S-band (index 2) is already functioning.

$$\forall l \in \mathcal{L}, b \in \{0,1\}, f \in F : A^l_{b,f} >= A^l_{b+1,f} \tag{4.13}$$

$$\forall l \in \mathcal{L}, b \in B, f \in \{0,1,2\} : A^l_{b,f} >= A^l_{b,f+1} \tag{4.14}$$

$$\forall l \in \mathcal{L}, f \in \{0,1,2\} : A^l_{2,f} >= A^l_{0,f+1} \tag{4.15}$$

When the demand uses a specific path, band, and fiber, all the links belonging to that path must be activated in the band and fiber chosen. The activation of a link ($A^l_{b,f}$) in a band and fiber is always 1 if and only if $X^d_{p,l}$ (1 if link l is included in path p) is 1 and $U_{p,b,f}$ (1 if demand d uses path p, band b and fiber f) is also 1, as defined in Eq. 4.16.

$$\forall d \in \mathcal{D}, c \in C, p \in P, l \in L, b \in B, f \in F : X^d_{p,l} \times U^d_{c,p,b,f} <= A^l_{b,f} \tag{4.16}$$

One demand can have one or more lightpaths, which can be grouped in one or more chunks. When in use, one chunk will always use one combination of path, band, and fiber. Since $U^d_{c,p,b,f}$ is 1 if a chunk of a demand d uses path p, band b and fiber f, the sum of all values from Eq. 4.17 should be less or equal to 1. Taking this into account, for a demand, the sum from Eq. 4.18 should be greater than or equal to 1, to guarantee that at least one chunk is being used.

$$\forall d \in \mathcal{D}, c \in C : \sum_{p \in P, b \in B, f \in F} U^d_{c,p,b,f} <= 1 \tag{4.17}$$

$$\forall d \in \mathcal{D} : \sum_{c \in C, p \in P, b \in B, f \in F} U^d_{c,p,b,f} >= 1 \tag{4.18}$$

In order to provision the requested traffic in each planning period p, all used chunks of lightpaths (that occupy a certain number of slots) must be placed in the spectrum. The sum of slots needed by all chunks in a link for a particular band and fiber must not exceed the

total stock of frequency slots in the link for the specific band and fiber. This is expressed in Eq. 4.19. Additionally, since all chunks have been placed for each demand, the sum of data-rate provided by them must be greater than the requested traffic for the demand at that year. This is expressed in Eq. 4.20.

$$\forall l \in \mathcal{L}, b \in B, f \in F : \sum_{d \in \mathcal{D}, c \in C, p \in P} Slots^{d}_{p,b} \times U^{d}_{c,p,b,f} \times X^{d}_{p,l} <= A^{l}_{b,f} \times Stock^{l}_{b,f} \quad (4.19)$$

$$\forall d \in \mathcal{D} : \sum_{c \in C, p \in P, b \in B, f \in F} U^{d}_{c,p,b,f} \times LP^{d}_{p,b} \times DR^{d}_{p,b} >= \tau_{d} \quad (4.20)$$

The following ILP is executed for each planning period and the results of the activated bands and fibers for each link in the network are recorded. It is important to note that the pre-processing calculates the worst-case reach of each BVT on each path and each band beforehand to avoid NLIN calculations during the band activation. This is done, since NLIN-based GSNR has been extensively studied in the previous chapters and is not the focus of this work.

The ILP is implemented using a Python-based Gurobi© optimizer and is run on an Intel Core i7 processor with 16 GB RAM.

4.3.3 Reinforcement Learning-based Upgrade Strategy

This section presents an upgrade strategy for band and fiber bundle activation based on Deep Q-Network (DQN)[4]. Unlike the ILP, which optimizes for all demands in a given planning period, this solution works on a per-demand basis in a given planning period. While there is a threat of achieving local optimal solutions instead of a globally optimal solution, the feature set size for an all-demand optimization is very large and compute-intensive. Therefore, the prediction of the upgrades to apply are iterated over each demand every planning period.

To start, Table 4.5 presents the notations, parameters, and variables used to define the RL solution. The RL solution will predict the action to take based on the variables, which are $Slots^{d}_{p,b,f}$, $Stock^{d}_{p,b,f}$ and $A^{d}_{p,b,f}$. The variable for slots ($Slots^{d}_{p,b,f}$) has information on how many slots are needed by a given demand. This quantity changes according to which path or band is chosen. Nevertheless, the index of fibers is added to keep the same size for all three variables. The stock variable ($Stock^{d}_{p,b,f}$) provides information on how many frequency slots are available in each path, band, and fiber. These values are pre-calculated for each path before the placement occurs. For this pre-calculation, $GeneralStock^{l}_{b,f}$ and $X^{d}_{p,l}$ are used.

[4]Refer to Appendix B for an overview of DQN.

Finally, the activation variable indicates if the K-shortest paths from demand at a respective band and fiber are active (already deployed) to use the resources. The values are extracted using $GeneralActivation^{l}_{b,f}$ and $X^{d}_{p,l}$.

Notation	Meaning
D	Array of demands
P	Array of paths
B	Array of bands
F	Array of fibers
d	Index of a demand
p	Index of a path
b	Index of a band
f	Index of a fiber
Parameter	**Meaning**
$X^{d}_{p,l}$	1 if path p from demand d includes link l, 0 otherwise
Variable	**Meaning**
$Slots^{d}_{p,b,f}$	Slots needed by a demand d in path p, band b and fiber f
$Stock^{d}_{p,b,f}$	Available slots for a demand in path p, band b and fiber f
$A^{d}_{p,b,f}$	For a demand d: 1 if band b in fiber f is active for path p, 0 otherwise
$GeneralStock^{l}_{b,f}$	Available slots for link l in band b and fiber f
$GeneralActivation^{l}_{b,f}$	1 if band b in fiber f is active for link l, 0 otherwise

Table 4.5: Definition of parameters and variables for the RL solution.

In RL, the agent performs an action, which causes changes in the environment, which is called a transition to a new state. In Deep-Q learning, the agent tries to find the best model to map a state to action with the highest reward. In Fig. 4.12, the four possible actions are presented, listed as follows:

1. Action 0 - no upgrade needed for a demand in a given planning period
2. Action 1 - one upgrade needed on the first shortest path for a demand in a given planning period (e.g. C to L or L to S on $p = 0$).
3. Action 2 - one upgrade needed on the second shortest path for a demand in a given planning period (e.g. C to L or L to S on $p = 1$).
4. Action 3 - one upgrade needed on the second shortest path for a demand in a given planning period (e.g. C to L or L to S on $p = 2$).

To create the RL model, a training environment is created. The training is based on a heuristic solution, which identifies the correct upgrade or no upgrade needed by each demand. The logic of the heuristic follows the below-described sequence:

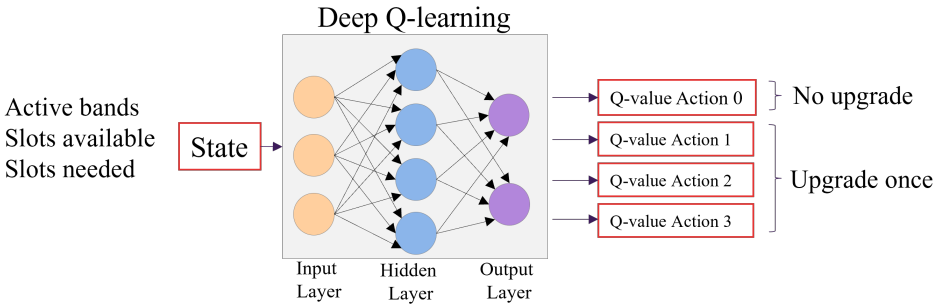


Figure 4.12: Deep-Q learning general architecture

1. If the number of slots needed ($Slots_{p,b,f}$) can be placed in active bands ($A_{p,b,f}$) of any of the three shortest paths, then there is no upgrade needed.

2. If the number of slots can only be placed after upgrading in 1 or 2 bands, then choose the path with the least active bands.

3. After each demand has identified the required action, the number of frequency slots available along all the links is updated before evaluating the next demand.

This way, there can be a guarantee of correct placement of the lightpaths and correct computation of $Stock_{p,b,f}$, which identifies the slots available for each path and at each band. For every correct decision taken by the RL model, a positive reward of +1 is given, whereas, for every wrong decision, a reward of -1 is given. The maximum reward is therefore the product of the total number of demands and the total number of planning periods. The pseudo-code and algorithm of the RL solution is provided in Algorithm 5 and Algorithm 6 (Appendix C).

As a first experiment, a model is trained over 1.11 million steps on a network with a single fiber-pair on all links and with the possibility of all three bands, i.e. C, L, and S . The maximum reward obtained is 2583 of 2860. When training, the mean reward per episode increases with the time steps and the loss progressively keeps decreasing till it saturates. In this case, the loss is $2e^{-3}$.

Subsequently, the RL solution is applied to a multi-fiber scenario with four fiber pairs, each with the possibility of having all three bands activated. The implementation of this solution is carried out in Python, using the Stable-Baselines 3 framework.

4.3.4 Multi-band Cost Models

To compare the different link upgrade strategies, the CAPEX for link upgrades to either a new band or a new fiber is also calculated. Taking into account the cost values of each component and the analysis with data from Table 4.6, the migration decision will comprehend the upgrade

from C, to C+L, to C+L+S-bands and finally to a new fiber. This is ordered from the lowest to highest cost. Assuming that lighting a dark fiber bundle is more expensive than an already operationg multi-band fiber by 0.6 C.U. [102], a priority is given to the S-band upgrade before the deployment of a new fiber to maximize the use of the spectrum.

Components	**C-band**	**L-band**	**S-band**
Amplifier (EDFA,EDFA,TDFA)	1	1.2	1.5
1x9 WSS	5	6	7.5
9x18 WSS	20	24	30
Transponder	36	43.2	54
MUX/DEMUX	0.04		
Fiber (per km)	0.8		

Table 4.6: Cost values relative to a C-band EDFA amplifier [102].

The scenarios considered for enabling and upgrading bands and fibers are as follows:

1. Enabling C-band: This provisioning includes the fiber deployment. The ILAs between each span are EDFAs operating in the C-band. At the ROADMs node locations, two additional EDFAs are placed as pre-amplifier and booster-amplifier. Also, a multiplexer and a demultiplexer are placed per ROADM degree as good practice, since this will later be used to deploy L-band and S-band ILAs. In addition, four Wavelength Selective Switchs (WSSs) that work in C-band are placed. The first two WSSs are placed per degree. The other two are placed to add and drop traffic from/to the BVTs. The number of BVTs will be equal to the number of lightpaths that are added/dropped at the ROADM node, and the quantity of WSSs will increase accordingly.

2. Enabling L-band: The upgrade to L-band at the ILA implies the addition of one EDFA in the L-band. At the ROADM nodes, two additional EDFAs operating in L-band are placed. Similarly, two WSSs for L-band are placed per degree and at least two WSSs for add/drop traffic to the BVTs. This number of WSSs will increase as the lightpaths that are added/dropped in the node increase.

3. Enabling S-band: This considers the addition of a TDFA at each ILA site. At the ROADM node, four WSSs in S-band are placed following the similar logic to that of the migration from C to L-band.

Upgrade	**Cost [C.U.]**
From C to L-band	$11.92 + 2.48 \times n_{span}$
From L to S-band	$15 + 3 \times n_{span}$
From S to C-band (new fiber)	$10.08 + 2 \times n_{span} + 1.6 \times d_{link}$

Table 4.7: Cost models of different band upgrades.

In Table 4.7, the cost model equations for each link upgrade are provided. In the table, n_{span} refers to the number of spans in the link and d_{link} refers to the total length of the link. The weights of the variables provided in Table 4.6. Additionally, the green-field scenario (status at the first planning period) assumes to have the first fiber pair deployed and the C-band activated on every link with no demands provisioned. Hence, the deployment of a new fiber pair is only applied when transitioning from S-band in the current fiber pair to the C-band in a new fiber pair.

4.3.5 Results and Discussions

This section presents the results of applying different methods for optical migration planning. The methods to be used are an ILP solution and an RL solution. The results obtained correspond to different networks with demands that reach the occupancy of four fiber pairs. The networks are Nobel-Germany, Sweden, and Nobel-EU.

From Table 4.8, it is evident that the ILP scales poorly with the increase in the network size and demands. In the case of the largest network under study, i.e. Nobel-EU, a 50% decrease in execution time for the RL-based solution is observed. This shows the advantage of using the RL-based solution for faster insights into multi-band upgrade strategies.

Network	ILP execution time (s)	RL execution time (s)
Nobel-Germany	284.24	161.67
Sweden	803.60	616.51
Nobel-EU	1275.61	691.39

Table 4.8: ILP and RL execution times for networks under study.

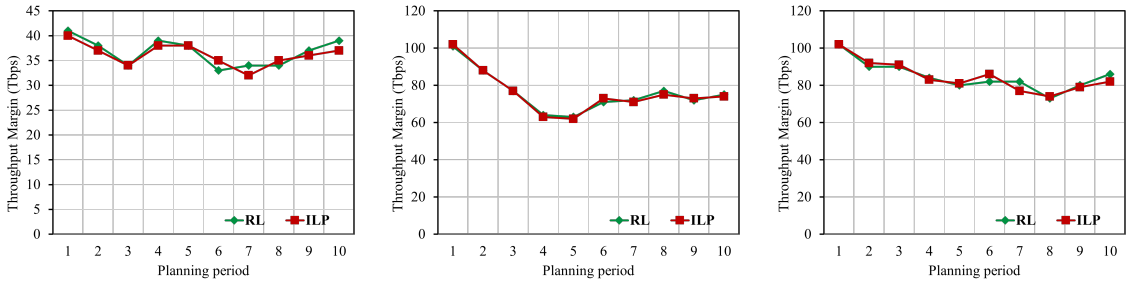
In Table 4.9, the testing metrics for the 4 fiber pairs are displayed. The accuracy of the Deep-Q learning model corresponds to the ratio of correct predictions divided by the total number of predictions. The lowest value is 90.92% for Nobel-Germany. The precision refers to the ratio between all the correct predictions of a class and the total predictions of this class. Higher precision leads to a higher recall score, which is defined as the ratio between all the correct predictions of a class to the total number of actual samples of this class. The weighted precision and recall are presented due to the inequality of the number of samples that refer to no upgrade with the ones of upgrade in one band. Finally, the F1 score is the harmonic mean between precision and recall, being defined by $F1\ score = 2.\frac{Recall.Precision}{Recall+Precision}$. This value gives the combined information of precision and recall and provides a better understanding of imbalanced classes.

The goal of each year is that the solutions provide all requested data-rate each year. The upgrades should be selected in a way that the status of the topology has enough spectrum

Network	Total Accuracy	Weighted Precision	Weighted Recall	F1 - score
Nobel-Germany	90.92%	99.11%	90.91%	94.83%
Sweden	96.05%	99.72%	96.05%	97.85%
Nobel-EU	95.12%	99.53%	95.12%	97.28%

Table 4.9: RL metrics.

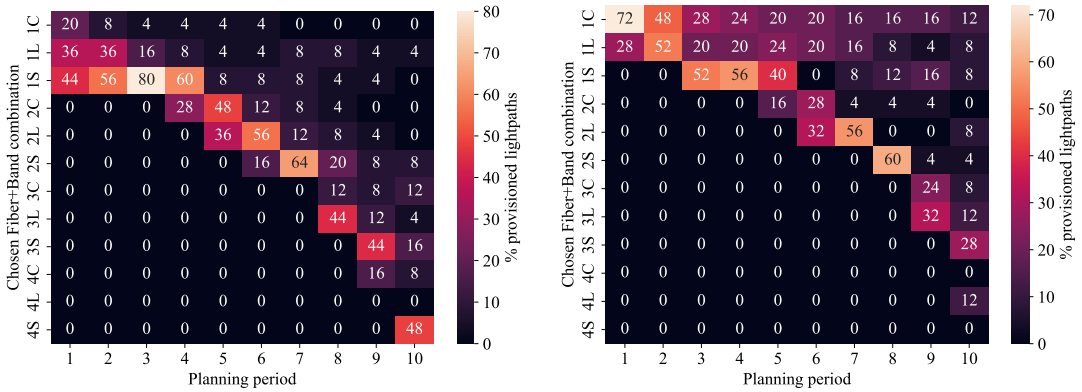
resources for the demands. Figs. 4.13a, 4.13b, and 4.13c display the accuracy of the data-rate placed by the ILP and RL solution.



(a) Nobel-Germany (b) Sweden (c) Nobel-EU

Figure 4.13: Difference between the cumulative provisioned data-rate and total aggregate traffic for Nobel-Germany, Sweden, and Nobel-EU networks.

As observed, the RL solution is not always accurate and can have demands not provisioned due to an erroneous decision. For example, if RL solution selects "Action 0", which represents no upgrade, when actually an upgrade was needed in order to fulfill the provisioning of the demand. Another case of a wrong decision is observed when the RL solution selects an upgrade in a path other than the needed one. Looking at the ILP solution, the difference to the data-rate is minimal. It is also observed that the ILP aims for upgrades one planning period earlier than the RL solution.



(a) RL solution (b) ILP solution

Figure 4.14: Candidate lightpath placement strategy comparison for Nobel-Germany network over 10 planning periods.

The upgrade strategy of ILP and RL solution differ from each other in terms of the upgrades undertaken. Fig. 4.14 shows the percentage split of the lightpaths placed in each band-fiber combination[5] for the Nobel-Germany network. The advantages of the global optimization per planning period are clear, with the ILP, which delays the usage of higher bands and fiber-pairs. On the other hand, the RL solution's agent seems to quickly activate higher bands and fibers.

However, despite the placement of lightpaths on higher bands in later planning periods (i.e. 3L to 4S), the locally optimal RL solution shows similar overprovisioning to the ILP solution, as shown in Fig. 4.15. In the initial years, the large overprovisioning is a design choice, because of the assumption that the deployed BVTs operate on the maximum possible data-rate for a particular's path reach (based on *HeCSON*). Adding additional planning periods can also lead to underprovisioning, if and only if all three bands on all four fiber pairs of a particular link cannot carry any additional traffic.

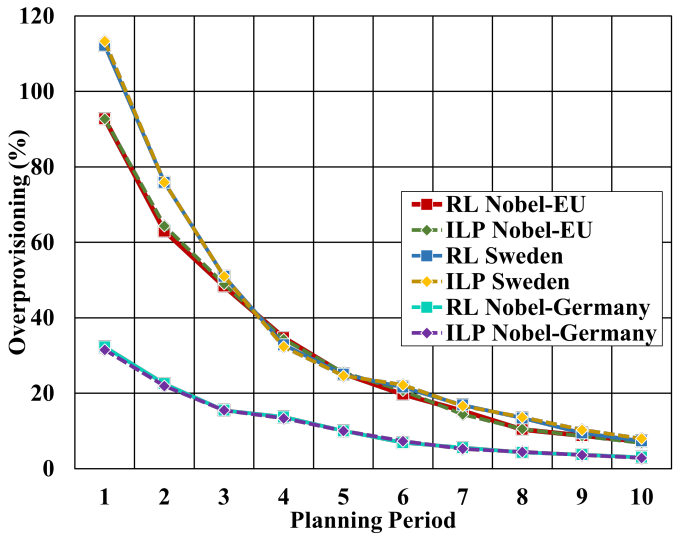


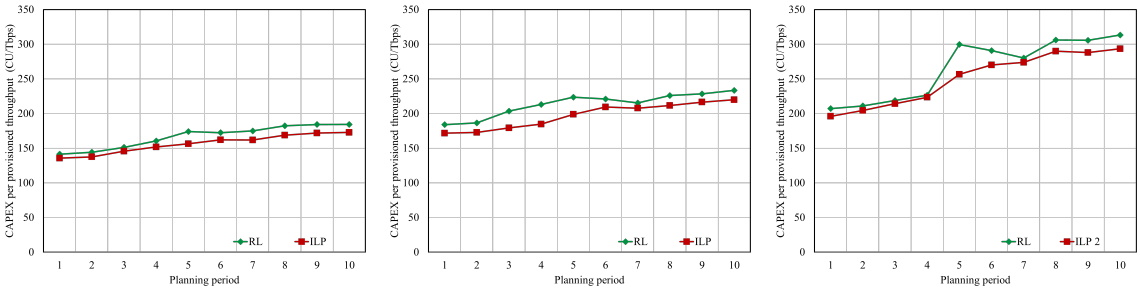
Figure 4.15: Overprovisioning for the three networks under study.

Despite the minimal shortcomings in traffic provisioning, the faster execution times make the RL solution as a viable tool for conducting migration feasibility studies for different traffic and network conditions which may arise in long-haul OTN.

A CAPEX study for the ROADM and BVT costs is also undertaken according to the cost model provided in Section 4.3.4. The first case involves the use of WSSs and BVTs for the three different bands. With the information of the architecture and the cost of the components, the cost-per-bit calculation is performed. Fig. 4.16 displays the cost per bit for each network comparing the ILP and RL solution. As expected, the RL solution for Sweden and Nobel-EU

[5]In Fig. 4.14, 1C refers to the C-band in the first fiber, whereas 4S refers to the S-band in the fourth fiber.

provides similar cumulative CAPEX per bit as compared to Nobel-Germany. An almost constant 30% overestimation of the cumulative CAPEX in the RL solution is observed, as compared to the ILP solution. This shows that on a network with lesser nodes and therefore lesser demand pairs, the RL solution offers lesser advantages as its agent has a lower recall score (ref. Table 4.9).



(a) Nobel-Germany (b) Sweden (c) Nobel-EU

Figure 4.16: Cumulative CAPEX per total throughput for networks under study across all planning periods.

It is also observed that the larger the network, the higher the cost-per-bit. For Sweden and Nobel-EU, there is a reduction in the cost-per-bit around planning period 7. In both cases, this is due to no additional fiber activation in that planning period, compared to the planning periods before and after.

4.3.6 Summary and Conclusions

Two solutions to plan band and fiber upgrades in long-haul OTNs are presented. The first solution is based on a multi-objective ILP, which tries to minimize the upgrades of the bands and fiber increments while meeting the requested traffic for all the demands in a given planning period. The second solution is a RL-based heuristic, which also tries to minimize the upgrades of the bands and fiber increments, for each demand individually in every planning year. Although the RL solution shows minimal disparities as compared to the optimal ILP solution, it must be noted that RL has up to 50% faster execution times. This is helpful if OTN operators need to filter through several permutations and combinations of input traffic and network scenarios. Post filtering, a more accurate migration planning can be done on the filtered scenarios using the ILP. Future work would involve improving the RL solution by aiming for a globally optimal reward function.

Chapter 5

Monitoring and Control of Disaggregated Optical Networks

In this chapter, **RQ3** is answered. As motivated in earlier chapters, BVTs are a constituent part of the long-haul flex-grid OTN architectures and are being deployed rapidly. However, post the planning and deployment phase, methodologies to monitor and control these devices need to be explored.

Specifically, encouraged by the disaggregation trend, several OTN operators are exploring new service models like OSaaS to increase their return on investments. In OSaaS, the OLS is owned and operated by the OTN operator, whereas the BVT is owned by the network end-customer (e.g. businesses, banks, government, etc.). As multiple end-customers use the same OLS, strict restrictions on the spectrum, modulation format, and launch power of each BVT must be enforced. Moreover, any failures on the OLS would not be visible to the end customers, who might want to switch their traffic to protection paths before service degradation occurs. For both these cases, a continuous monitoring solution is needed. The latter part is also interesting for a failure detection use case, since the continuous data collection may provide insights into the OLS health, thereby allowing end customers to protect their traffic from gradual QoT degradation. In this chapter, we investigate methods to allow OSaaS users to detect soft failures of their BVTs as soon as possible while using data collected from a live production network.

The second aspect of controlling disaggregated OTNs arises from the availability of open-source platforms to build an end-to-end solution to control disaggregated OTNs. These fully disaggregated devices aim to bring down network upgrade costs by allowing multi-vendor interoperability. However, before such a network is deployed in the field, several control plane functionalities need to be developed and tested in simulation environments. To enable this, the latter half of this chapter discusses in detail a proof-of-concept of a cloud-

based disaggregated OTN simulation, which is controlled using open-source SDN controllers. Moreover, the planning, orchestration, and restoration of optical services can be decoupled from the simulated network, by using Transport Layer Security (TLS)-enabled north-bound REST APIs exposed by OpenDayLight TransportPCE, which is an open-source optical domain controller (see Section 2.4).

5.1 Monitoring failures for OSaaS users

OTNs allow network operators to efficiently adjust the optical bandwidth of each BVT in the network, based on the demand's bit rate and transmission distance [103]. However, despite their advantages, flex-grid OTNs are more susceptible to soft failures than fixed-grid OTNs due to the increased complexity of the individual network components and various link impairments [104]. Soft failures are defined as those failures that affect the signal quality through slowly varying phenomena that manifest themselves as anomalies in the Optical Performance Monitoring (OPM) data [105]. These anomalies can gradually raise the Bit Error Rate (BER), leading to potential service downtime. Therefore, detecting, identifying, and counteracting these soft failures can significantly reduce system disruptions and repair costs.

To leverage the full potential of the infrastructure of OTNs through flexible resource utilization, the concept of OSaaS has emerged [106]. In OSaaS, BVTs are owned and controlled by the service end user, while the OLS, which provides signal equalization, transportation, and amplification, is controlled by the OTN operator. This method enables flexible resource utilization inside the dedicated customer spectrum of the OSaaS and eliminates unnecessary O-E-O conversions [107].

In light of the anticipated high demand for OSaaS, this study addresses the question of how OSaaS users can detect and identify failures at an early stage while only using information that is available at the end-user BVTs. The approach is completely independent of OLS element telemetry data, as OLS-related parameters and configurations are not known to the OSaaS service user and will thus not be considered in the scope of this study.

In this study, OPM data from five commercial BVTs deployed on a live long-haul pan-European network is extracted. Acting as an OSaaS user with five BVTs, no information of the underlying network is possessed. To enable single-ended measurements from the test site, a physical optical loop-back is featured in the far-end ROADM in the network, effectively doubling the optical transmission distance. The total transmission distance of the CUTs is 1792 km and they are continuously monitored for 45 days. The input data points are considered to be the OPM data of all the CUTs.

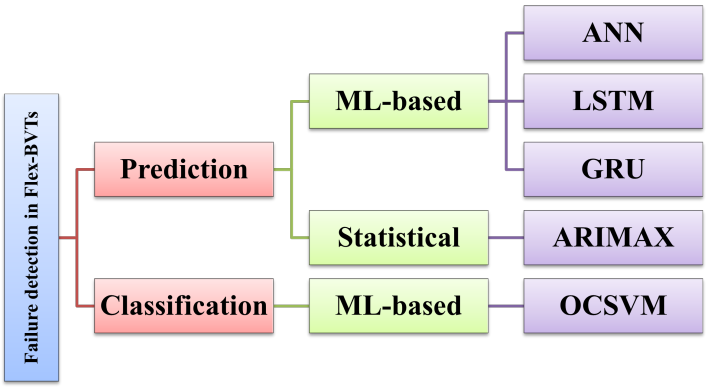


Figure 5.1: Failure detection algorithms investigated in this work.

In long-haul OTNs, soft failures resulting in the gradual degradation of the signals occur infrequently compared to hard failures, introduced by fiber cuts, power outages, fatal equipment failure, or human errors in handling fibers. Thus, soft failure detection can be conceptualized as an exercise in anomaly detection. A common method in anomaly detection is to predict the expected value of a parameter of interest and compare it to the monitored value. If the difference between the prediction and the monitored value exceeds a certain threshold, it is defined as a failure.

As shown in Fig. 5.1, failure detection can be broadly divided into two categories: detecting failures based on the prediction of certain input parameters, such as Optical Receive Power (ORP) or BER; and detecting failures based on classification algorithms. Due to the large amount of time-series data generated by continuous monitoring, ML-based algorithms are well-suited for predicting or classifying failures. To compare the benefits of using ML-based algorithms for ORP prediction, a statistical-based algorithm, namely, Autoregressive Integrated Moving Average with Explanatory Variable (ARIMAX) is also implemented. Additionally, the One Class Support Vector Machine (OCSVM) is implemented as a classification-based approach.

In this work, first, a comparison of the prediction-based ML algorithms (Artificial Neural Network (ANN), Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU)) and the statistical-based ARIMAX algorithm is carried out. Then the best-performing prediction-based ML algorithm is compared with the classification-based OCSVM in terms of accuracy of prediction/classification, misclassification rate, and time taken for predicting or classifying a failure.

5.1.1 State-of-the-Art Analysis

In recent years, OSaaS margins and QoT have been experimentally characterized using several manual network testing techniques [107]. The results of such studies enable network operators to operate the BVTs in a low-margin regime. However, the amount of data that can be collected is limited to a few devices, and finding patterns manually becomes increasingly difficult.

In optical network failure detection, several methods have been analyzed, including the use of state vector machines for detecting and identifying filter shift and tight filtering [108]. The residual-based SVM was found to perform best in terms of accuracy and robustness. Additionally, a new method of analyzing constellation diagram images using a Convolutional Neural Network and the Density-Based Spatial Clustering of Applications with Noise algorithm was presented [109]. However, these methods are reported to be computationally intensive, causing failure detection delays.

A method for detecting, localizing, and identifying potential faults using SDN integrated knowledge was also proposed [110]. The failure detection uses optical power level abnormalities and the localization works by network topology mapping. However, questions remain regarding whether proposed solutions also work on real field data, and which ML techniques are best suited for this purpose. Furthermore, there are currently no studies evaluating failure detection for OSaaS users, where the only knowledge of the network is available from OPM data at the end-user BVTs.

5.1.2 Measurement Setup and Data Collection

To collect data for training and evaluation of the ML models, a commercial pan-European live network owned and operated by Tele2 Estonia spanning 2869 km is used (map in Fig. 5.2). Besides the live channels carrying production traffic, five ADVA TeraFlex™ BVTs [18] were installed to generate data for testing. To create training data for the underlying study, five test channels were inserted into a dedicated add/drop port of the ROADM in the test site using an 8-port splitter/combiner module. Then, a 400 GHz wide OSaaS with a central frequency of 193.95 THz within the OLS C-band was configured in all traversed ROADMs.

A physical loopback was featured in the far-end ROADM location, to enable single-ended measurements from the test site and create a link length of 1792 km. BVT configurations 200 Gbps QPSK 69 GBaud and 200 Gbps DP-16QAM 34 GBaud were used in data collection, maintaining the nominal power spectral density of the network, when switched between configurations.

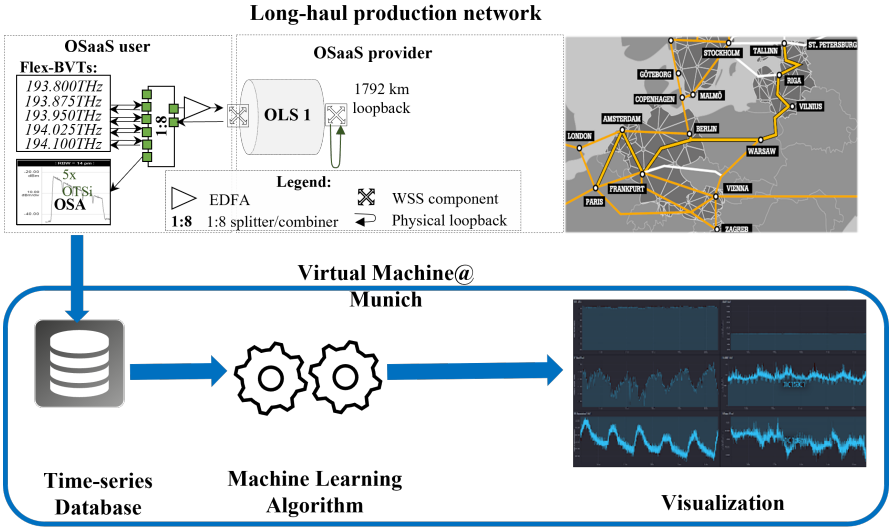


Figure 5.2: Overview of the production network and the automated data collection framework.

5.1.2.1 Collecting OPM Data

Fig. 5.2 depicts an overview of the data collection framework, using a Python-based script to extract monitoring telemetry via NETCONF requests, polling the devices every 30 seconds, and then storing the returned information in a time-series database. From each CUT, a total of 7 OPM parameters were extracted, namely, Carrier Frequency Offset (CFO), Chromatic Dispersion Compensation (CDC), Differential Group Delay (DGD), Pre-Forward Error Correction Bit Error Rate (pre-FEC BER), ORP, OSNR, Q-factor, and the electrical SNR[1].

5.1.2.2 Generating Failure Data

Two types of soft failures are artificially introduced on the CUTs in our analysis, namely, power degradation and inter-channel interference. To simulate gradual power degradation, the transmit power of each CUTs was reduced by 0.2, 0.5, and 1 dBm every minute until the signal cannot be decoded. While the ORP of each CUTs was gradually decreasing, the pre-FEC BER also increased. Henceforth, the power degradation failure is referred to as **Failure 1**.

Inter-channel interference occurs when two neighboring channels are too close to each other. This failure was artificially introduced by shifting the tuned frequency of the CUTs first to the left and then to the right by 6.25 GHz frequency steps, reducing the channel spacing to the left and the right neighboring channel, respectively. It can be seen that the BER of the overlapping CUTs is increasing. This failure is henceforth referred to as **Failure 2**.

[1]It is important to note that the OSNR reported by the BVT is actually the calculated GSNR since optical DSPs cannot conduct measurements. The electrical SNR is obtained after the optical DSP post-processing and error-correction.

5.1.3 Methodology

As discussed earlier, soft failures occur rarely and it is difficult to build a large enough failure data set to train a supervised ML model on data including failures. Hence, the failure detection part of this work is considered an exercise in anomaly detection using supervised and semi-supervised learning methods based on data without failures.

5.1.3.1 Data Analysis

The OPM data used for training and hyperparameter optimization covers the period from 18th November 2020 until 2nd January 2021 and thus spans 45 days. Further, the OPM data within this time period was stable with minimal human interventions and hard failures. An additional 4.5 days of data were collected and used as the testing data for the trained models.

For analyzing the OPM data from all CUTs, correlation methods are used. These methods measure the relationship between each OPM parameter and provide non-correlated features as input to the ML models. Pearson Correlation, Spearman Index, and Maximum Information Co-efficient techniques are used to remove correlated parameters. From Pearson correlation coefficients, it is observed that the strongest positive linear correlations are between the Q-factor, electrical SNR, and OSNR. These three variables behave very similarly in terms of linearity. The ORP follows similar patterns as the electrical SNR but is weakly correlated. Looking at negative correlations, it is noticeable that pre-FEC BER is strongly negative linear dependent on the four previously mentioned variables. From the Spearman ranks, the existence of a non-linear correlation between ORP and pre-FEC BER can be inferred. Finally, the results of Maximum Information Correlation suggest a weak correlation of CDC with the strongly correlated parameters of the electrical SNR, OSNR, pre-FEC BER, and ORP.

Although the seasonality and effect of ambient temperature on various OPM parameters were also studied, they are not included in the scope of this work since the equipment is always operating in temperature-controlled spaces.

5.1.3.2 Feature Scaling and Selection

As all parameters monitored have different ranges, the data is normalized to ensure unitless input for the ML algorithms. This is necessary since most algorithms expect scaled input data. For failure detection, the OPM data is transformed such that each feature individually ranges from zero to one instead of from its minimum to maximum. Every entry x_i of each feature vector $\mathbf{x}$ is converted to [111]:

$$x_{i,MinMaxScaled} = \frac{x_i - min(\mathbf{x})}{max(\mathbf{x}) - min(\mathbf{x})} \tag{5.1}$$

Choosing the right features as input plays an important role in terms of accuracy. Fewer parameters not only lead to decreased computational time and complexity but also better interpretability. Of all 7 extracted OPM parameters (ref. Sec. 5.1.2.1), only pre-FEC BER and ORP are selected as input features. This is justified based on a combination of data analysis and existing literature [105]. Other parameters such as pre-FEC BER, OSNR, Q-factor, and electrical SNR provide information to augment ORP. However, since the correlation analysis shows that they all carry similar information, a reliable and up-to-date pre-FEC BER can be selected.

5.1.3.3 Hyperparamater Optimization

As shown in Fig. 5.1, failure detection can be formulated either as an ORP prediction problem or as a classification problem. For ORP prediction, ARIMAX algorithm is used as a statistical baseline to compare ML algorithms. For ARIMAX, hyperparameters p, q, and d have to be chosen [112]. A grid search is performed to iteratively find optimal values of the hyperparameters. The optimal set of parameters is defined by the set with the lowest Akaike information criteria value [113]. It gives a relative measure of whether the model is a good fit for the data, also taking into account the model's complexity. The results of the grid search indicate that a combination of low values of p and d as well as a high value for q result in the lowest Akaike information criteria value. Table 5.1 shows the values chosen for ARIMAX, given the collected data.

Model	Hyperparameter	Value
ARIMAX	p	0
	d	1
	q	4
ANN, LSTM, GRU	Hidden Layers	3
	Activation function (ANN only)	ReLU
	Optimizer	Adam
	Batch size	50
	Epochs	100
	Test split	0.25
	Validation split	0.2
OCSVM	Kernel	4th degree polynomial
	ν	0.0005
	k	5

Table 5.1: Hyperparameters for ARIMAX, ANN, and OCSVM models.

Since training an ANN takes longer compared to other algorithms and large hyperparameter space, a combination of knowledge about the general behavior of ANNs and randomized search is used to find the optimal hyperparameters. Randomized search evaluates a given number of random combinations by selecting a random value for every hyperparameter at

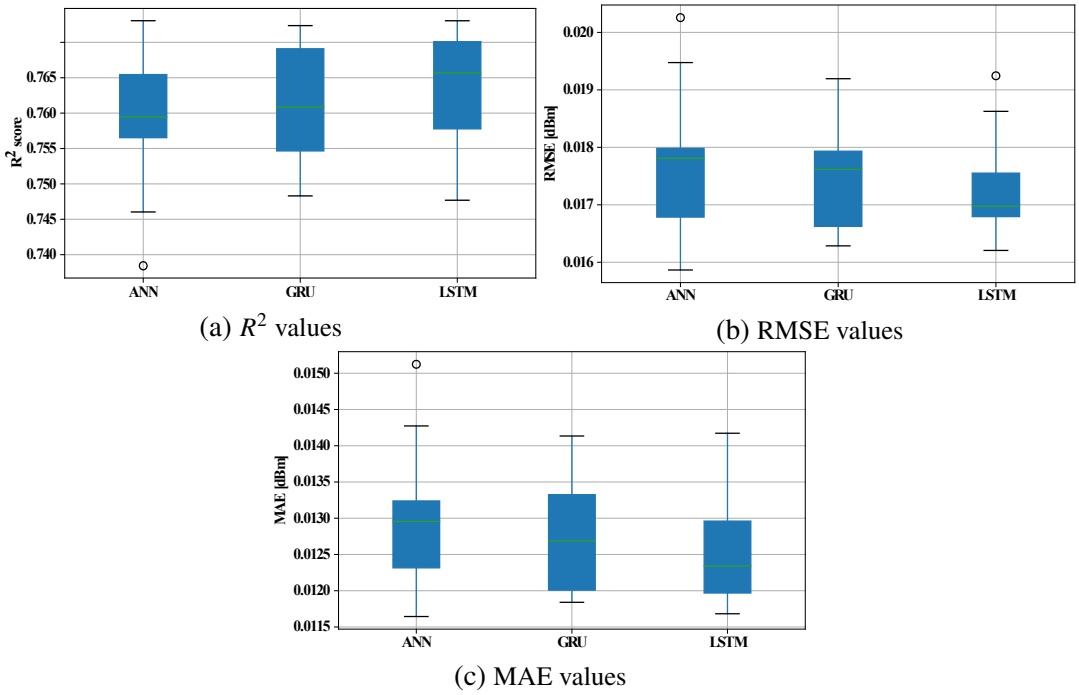
each iteration. Therefore, this search method can find better models within a limited evaluation time by effectively searching a larger configuration space [114]. The architecture of the ANN needs to be pre-defined and cannot be set by a randomized search. Deeper networks containing more hidden layers have a much higher parameter efficiency, resulting in fewer neurons and thus less computation time for training. Therefore, three hidden layers in addition to a flattening and a dropout layer are chosen. The flattening layer undertakes dimensional reduction of the time series and the dropout layer avoids model over-fitting. After a 33-run randomized search, an R^2 score of 0.9955 is achieved. It is assumed that the same hyperparameters (excluding the activation function) are ideal for the LSTM and the GRU models. The train-test split for the data is selected as 80%-20% and data are shuffled before splitting to ensure homogeneity.

To find the best values for the hyperparameters of the OCSVM, a grid search with the validation split parameter k (i.e. number of datasets of equal size), the ratio of outliers ν, and kernel function is performed. Validation split is used to assess how a model can predict or classify new data of an independent data set not used for estimating it, in order to reduce overfitting and bias during the training. From the results it becomes clear, that kernel functions with the best performance are polynomial functions with a degree of two and four. They both perform with 100% accuracy (no misclassified data points) for small ν. Therefore, the hyperparameters for OCSVM are selected as: kernel = polynomial with degree = 4, $\nu = 0.05\%$, and k = 5.

Once the ML models were trained, additional data was extracted from the production network for another 4.5 days. This “unseen" data, containing 13000 datapoints were injected with 6 artificially introduced failures (four of type Failure 1 and two of type Failure 2). Each trained ML model is run independently on this test data while predicting the ORP at each time step.

5.1.3.4 Dynamic Threshold Calculation

The ANN model is further augmented by a dynamic threshold calculation to reduce the prediction error [112]. From the prediction error of historical values, a single-dimensional vector $\mathbf{e}$ is created and each value is further smoothed using an exponentially-weighted moving average in order to obtain $\mathbf{e}_s$. A set of candidate thresholds is computed as $\epsilon = \mu(\mathbf{e}_s) + \mathbf{z}\sigma(\mathbf{e}_s)$, where $\mathbf{z}$ is an ordered vector of the k highest deviations of $\mathbf{e}_s$ above the mean $\mu(\mathbf{e}_s)$. k is a tunable hyperparameter between 2-10. Therefore, the actual threshold can be selected as



(a) R^2 values

(b) RMSE values

(c) MAE values

Figure 5.3: Analysis of different ML algorithms for ORP prediction method.

shown:

$$t = \frac{\frac{\Delta\mu(\mathbf{e}_s)}{\mu(\mathbf{e}_s)} + \frac{\Delta\sigma(\mathbf{e}_s)}{\sigma(\mathbf{e}_s)}^2}{|\mathbf{e_a}| + |\mathbf{E}_{seq}|} \tag{5.2}$$

$$\text{where} \ : \Delta\mu(\mathbf{e}_s) = \mu(\mathbf{e}_s) - \mu(\{e_s \in \mathbf{e}_s | e_s < \epsilon\}), \tag{5.3}$$

$$\Delta\sigma(\mathbf{e}_s) = \sigma(\mathbf{e}_s) - \sigma(\{e_s \in \mathbf{e}_s | e_s < \epsilon\}), \tag{5.4}$$

$$\mathbf{e}_a = \{e_s \in \mathbf{e}_s | e_s < \epsilon\} \tag{5.5}$$

$$\mathbf{E}_{seq} = \text{continuous sequences of } e_a \in \mathbf{e}_a \tag{5.6}$$

The error threshold for ORP prediction with ANN is calculated one timestep ahead, thereby allowing reducing the overall prediction errors and improving the accuracy.

5.1.4 Results and Discussions

Once the hyperparameters of the ML models have been chosen, the ML algorithms are compared using two methods. First, the R^2, Root Mean Square Error (RMSE), and Mean Absolute Error (MAE) values of the ORP prediction-based ML models are compared. Then, the three models are compared with ARIMAX in terms of computation time. Finally, the best-performing prediction-based ML method is compared with the classification-based OCSVM method in order to detect artificially introduced failures in the network.

To evaluate the accuracy of the ORP prediction approaches, R^2 score, RMSE, and MAE are computed with 10-fold cross validation. The results shown in Fig. 5.3 are based on the mean of ten different sets of unseen test data.

As seen in Fig. 5.3a, the R^2 score is highest for the LSTM model, with a 0.005 increase in the median as compared to ANN and GRU. From Fig. 5.3b it becomes clear, that again LSTM performs better compared to the other two algorithms. Fig. 5.3c shows, that the LSTM model has the lowest MAE whereas The GRU outperforms the fully connected ANN by about 0.005 dBm. In summary, LSTM outperforms the other two neural networks slightly for all three accuracy metrics. The GRU and the fully connected ANN are very similar in the context of R^2 score and RMSE, but GRU has a slightly lower MAE.

Due to the fact that the ARIMAX model has very a high computation time for large datasets, which is shown in the next section, it can not be applied directly to the entire unseen data. Nevertheless, a piecewise comparison using five different datasets each of 1000 points, has revealed 17%, 15%, and 14% higher RMSE values than ANN, LSTM, and GRU, respectively. Furthermore, the MAE of ARIMAX is 16%, 14%, and 13% higher compared to ANN, LSTM, and GRU, respectively.

Fig. 5.4 shows, how the calculation time for each model increases with increasing amounts of points to predict. While ARIMAX quickly escalates its calculation time in order of minutes (right y-axis), GRU and LSTM stay in the range of 1-12 seconds for 5000 data points (left y-axis).

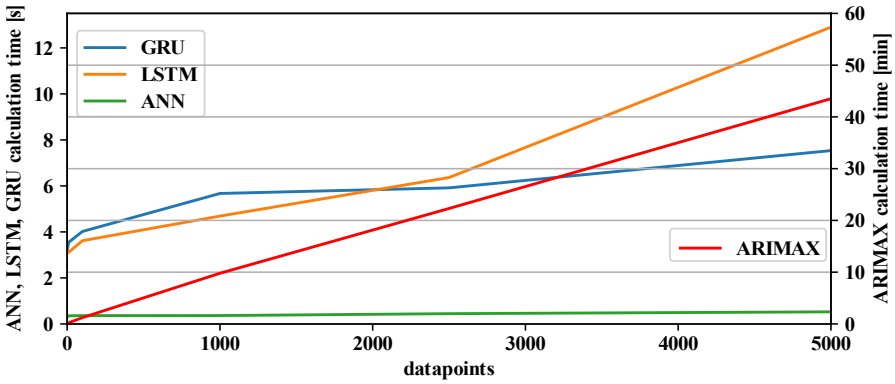


Figure 5.4: Computation time for ORP prediction algorithms.

In terms of failure detection, the ANN model in combination with the dynamic threshold is compared to OCSVM. To calculate the detection accuracy, six artificially introduced CUTs failures (ref. Sec. 5.1.2.2) as well as 13000 unseen faultless data points are used.

Fig. 5.5 the total duration of each failure in comparison with the time until it is detected by each algorithm. For failure 1, only three out of six errors are detected by OCSVM. This is because a correctly classified failure is defined as a binary operation between *failure* and

no failure. Further, for three measurements of Failure 1, the ORP and pre-FEC BER values before and after the failures occurred are misclassified as failures by OCSVM and hence no failure is detected, when an actual failure occurs. All other failures are correctly detected by both algorithms. For ANN with a dynamic threshold, 2 out of 13000 (< 0.01%) datapoints are misclassified as failures indicated by an ORP prediction error lying above the dynamic threshold. OCSVM, misclassified 112 out of 13000 datapoints (approx. 0.6%), thereby leading to a higher number of false positives. The low performances of OCSVM can be explained by the fact that the model is adapted to the training data. Therefore, the OCSVM algorithm is not capable of adapting to new ORP ranges and trends, which is the case for the real field ORP data.

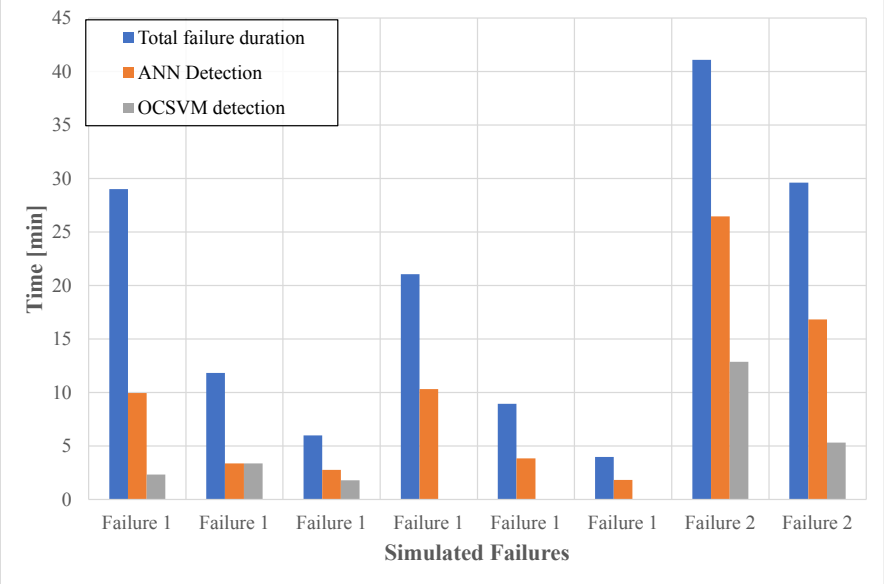


Figure 5.5: Computation for times for various failure detections.

The mean of the difference between failure start to detection divided by the total failure duration is 24.29% for OCSVM and 47.54% for the ANN with a dynamic threshold approach. Hence, the ANN approach performs sufficiently well as well, detecting the failure almost always within the first half of the total failure duration. The computation time is 0.0057 seconds for OCSVM and 0.7262 seconds for the detection with ANN including the dynamic threshold calculation.

5.1.5 Summary and Conclusions

In this work, several EON fault detection and identification algorithms were compared and evaluated based only on the OPM data available at the end user's transceiver. Overall, five algorithms were evaluated on OPM data from a 1792 km loopback link from a live production network, running for 45 days. Furthermore, two different types of failures were recreated to evaluate the models. For the failure detection, ORP values were predicted with three different neural network architectures as well as with the statistical approach ARIMAX. Based on the resulting prediction error, a dynamic threshold was used to augment the ANN's decision.

Furthermore, as a second approach for failure detection, an OCSVM was built based on faultless data. The evaluation showed, that although LSTM has the highest accuracy amongst the ORP prediction algorithm, the fully connected ANN has the shortest calculation time. Comparing the ANN model augmented by a dynamic threshold calculation with OCSVM, failures are detected earlier with OCSVM, but with a higher accuracy using the ANN approach. While the ANN approach has a 100% detection rate and 99.989% accuracy rate for failures and faultless data, respectively, the OCSVM achieved a 62.5% failure detection rate and a 99.45% accuracy rate.

Nevertheless, as external conditions might change over time, it is recommended to retrain the prediction ANN as soon as the standard deviation of one day exceeds 1 dBm. To overcome problems with false failure detection, a failure should be recognized only if two consecutive failures have been detected by the algorithm. This would however delay the time of detection by 30 seconds.

Through this work, it is shown that certain degradation in the network can be detected quickly without having complete knowledge of the underlying OLS. This can help in quick service restoration and root cause analyses. Machine learning algorithms can benefit from the augmentation of non-ML-based algorithms to improve the accuracy of prediction. Future work includes failure identification for OSaaS users using additional input parameters.

5.2 Open-Source Control of Open Disaggregated Optical Transport Networks

Network operators across the world are planning 5G deployment for industry automation and mission-critical application support for enterprises. DWDM based OTN is one of the major enablers for these technologies, due to its ability to carry high-volume high-speed data over long distances[115]. With the increase in software capabilities, vendor lock-ins for network and device management are bound to occur. In such a network, vendors typically use proprietary data models and interfaces, thereby restricting device choice. To overcome vendor lock-ins, fully Disaggregated Optical Transport Network (DOTN) are proposed, consisting of Open Terminals and OLS [116].

To enable interoperability among vendors and to drive down deployment expenditures, MSAs have been set up to define standardized data models and interfaces for optical networking devices. To this end, OpenROADM MSA is one of the industry-led standards, which defines required data models for devices, networks, and services [51]. Using such models, underlying optical network devices like ROADMs and Optical Transponders (XPDRs) can be configured and monitored.

To be compliant with OpenROADM MSA, vendors need to implement and expose standardized APIs, which can be used by open-source network management and controllers to control such devices. An optical domain controller like OpenDayLight TransportPCE [117] uses these data models to expose south-bound interfaces for control of devices, and north-bound interfaces for communication with applications[2].

Before deploying an open-source disaggregated controller and orchestrator on a live network, network operators need to evaluate it in a “sandbox" of simulated optical devices. Such a setup helps network operators not only to execute “what-if" scenarios but also to simulate rare device outages and understand how the control plane reacts to them. Additionally, interoperability demonstrations are typically restricted to a few network devices, thereby creating a need for a large-scale network simulation to understand network-wide effects.

Therefore, in this work, a cloud-based optical network simulation is presented. This cloud-based simulation can communicate with an open-source optical domain controller. Further, a fully operational network planner and optical service orchestrator are demonstrated, which communicates with the domain controller using secure RESTCONF APIs. Finally, the network planner and service orchestrator is offered as a web-application (WebApp), which can be deployed and accessed remotely.

[2]This concept has been introduced in Fig. 1.5 (Chapter 1

5.2.1 State-of-the-Art Analysis

Traditionally, lightpaths were deployed manually by network operation control technicians, by accessing each of the required devices individually; adding the configuration settings, and switching on the laser using CLI commands. However, there has been a recent shift towards the implementation of SDN concepts to DWDM-based OTN [118].

For control of simulated optical networks, Troia *et. al.* presented an SDN based routing demonstration [119]. However, no information was provided on the usage of standardized data models, or on the QoT estimation of the deployed lightpaths. Recently, Vilalta *et. al.* presented a concept of an optical network digital twin, which envisions the usage of emulated devices for a metro optical network [120]. This work highlighted the capabilities of open-source technologies in future optical network management. However, network-wide studies were not discussed in this early stage of proof-of-concept.

Once lightpaths are deployed and are "in-operation", they need to be constantly monitored to check important parameters like the GSNR, launch power, bit error rate, etc. This is done by pulling OPM data from devices. Several optical domain controllers rely on NETCONF and SNMP protocols to retrieve OPM data. However, these solutions are often not scalable. A workaround is the use of streaming telemetry based on gNMI protocol as defined by OpenConfig YANG models [121]. Such streaming telemetry mechanisms have been experimentally validated on actual optical networking elements, and have proven to be more effective than other methodologies [122].

Once the network monitoring is in place, in order to effectively close the loop, a service re-routing strategy for failure protection needs to be implemented. Several works focus on creating algorithms in order to effectively restore services, with minimal loss of service. The push-pull algorithm reconfigures lightpaths in the network without traffic disruptions while considering only 100 Gbps QPSK configuration in their coherent network deployment [123]. However, such a technique is not validated in an heterogeneous flex-grid OLS deployment. Therefore, there is a need for heuristics that restore lightpaths in a highly volatile elastic optical network environment.

Although many works exist in individual parts of optical network planning, monitoring, and restoration, however, these studies offer individualistic solutions and ignore network-wide or system-wide planning and deployment intricacies. Hence, there arises a need to cover these three important aspects of the optical network and propose a single solution that allows network operators to simulate complex edge-cases and prepare their deployed networks for any eventuality.

5.2.2 In-operation Control Mechanisms

Once a network is planned and deployed, each component needs to be monitored, so that the services can be re-routed in case of a component or a link failure. To monitor and control optical networks, the principle of SDN is applied. SDN allows the separation of the control plane of the network from the data plane and offers multiple interfaces which allow network users to communicate with the devices. To control optical networking elements, an optical domain controller is required. This controller connects to underlying optical devices using NETCONF protocol [124]. Such a connection from the controller to the network elements is also known as a south-bound interface. Similarly, the controller also exposes several north-bound application programming interfaces like RESTCONF [125]. Further, these RESTCONF commands can be automated to create, delete, and manage optical services using a multi-domain network orchestrator.

Once the network control is set up, several algorithms can be added to the network orchestrator level to enable the re-routing of optical services in case of failure or loss of connectivity from the connected nodes. The focus, therefore, should be to develop algorithms that can take as an input real-time network information and translate it to RESTCONF-based commands to configure the underlying optical networking elements. In this work, a proof-of-concept to implement such a novel system is undertaken. Such a proof-of-concept allows optical network researchers to study software design, understand the time taken by the control plane to configure devices, and resolve any scalability issues, in order to provide a holistic network control solution.

5.2.3 Demonstration Setup

As seen in Fig. 5.6, the Nobel-Germany DOTN is set up on a cloud server hosted in Open SDN & NFV Lab, Berlin (OSNL), Berlin, Germany. Each network element uses the Network Topology Simulator (NTS) framework that simulates OpenROADM devices (ROADMs and XPDRs) as docker containers with a NETCONF server running in it. For setting up the 17-node Nobel-Germany DOTN topology, 17 ROADM network elements and 52 XPDR elements are deployed using deployment scripts, which are available in an open-source repository [126]. The number of XPDRs at each node corresponds to its node degree, thereby allowing the formation of at least 26 end-to-end lightpaths in the network.

For service planning and deployment, a Java-based web application called *HeCSON* WebApp is used [3]. The *HeCSON* WebApp will be presented locally at the conference. Apart from an integrated DOTN planning, *HeCSON* can also connect and send service requests to the TransportPCE controller using a Virtual Private Network (VPN) tunnel. To reconfigure

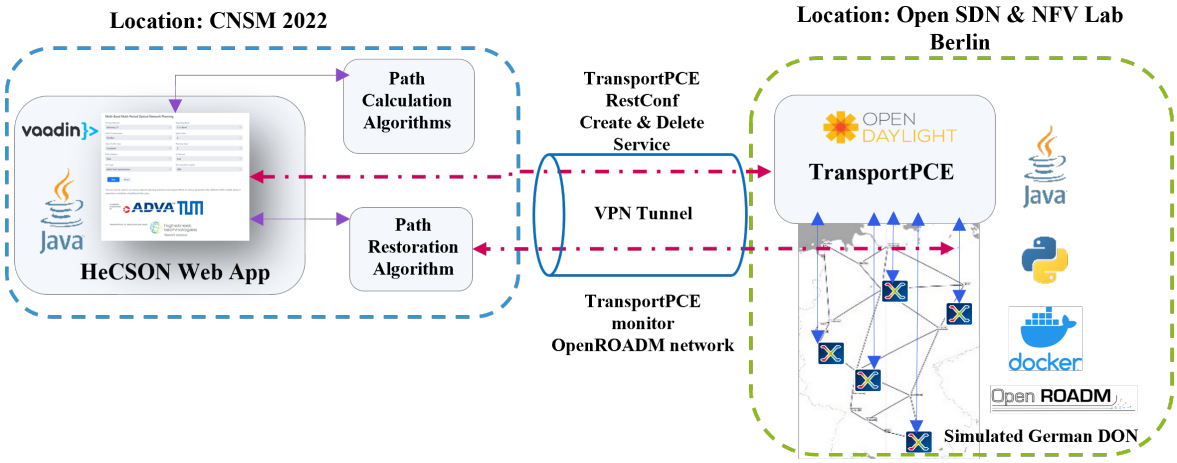


Figure 5.6: Demonstration setup and system architecture.

optical services, *HeCSON* also includes a polling-based network element monitoring. All the REST API commands used by *HeCSON* (*Service Create, Delete, and Monitor Network*) are sent to TransportPCE using TLS encryption. Fig. 5.6 visualizes the complete demonstration setup.

5.2.4 Use cases and Interaction

Two use-cases are demonstrated with this setup, namely, *a)* planning and provisioning of services onto a simulated Nobel-Germany DOTN (use-case 1), and *b)* monitoring failures and restoring services (use-case 2). The prerequisite for both the use-cases is that The TransportPCE and the simulated Nobel-Germany DOTN have to be successfully configured and deployed.

In use-case 1, users have access to *HeCSON* WebApp, through which a network planning of the simulated network can be run. For each candidate lightpath which has a valid QoT, clicking on the “Add to TransportPCE" button triggers an Hypertext Transfer Protocol (Secure) (HTTPS) POST request from *HeCSON* to TransportPCE deployed in OSNL. Once TransportPCE configures the required network elements like ROADMs and XPDRs in the Nobel-Germany DOTN, the administrative state of the service is set to “inService". This shows successful end-to-end provisioning of the service. The attendees can now view and interact with the deployed service on the network map hosted in OSNL, as shown in Fig. 5.7.

For use-case 2, users can start network monitoring via the *HeCSON* WebApp. *HeCSON* retrieves the current state of all the connected network elements, and checks for a change in their status. In case of a loss of connectivity, it first triggers an HTTPS POST request to delete the affected services and then create new services using the unaffected network element. When all affected services are restored, the user receives a “Services Restored" notification.

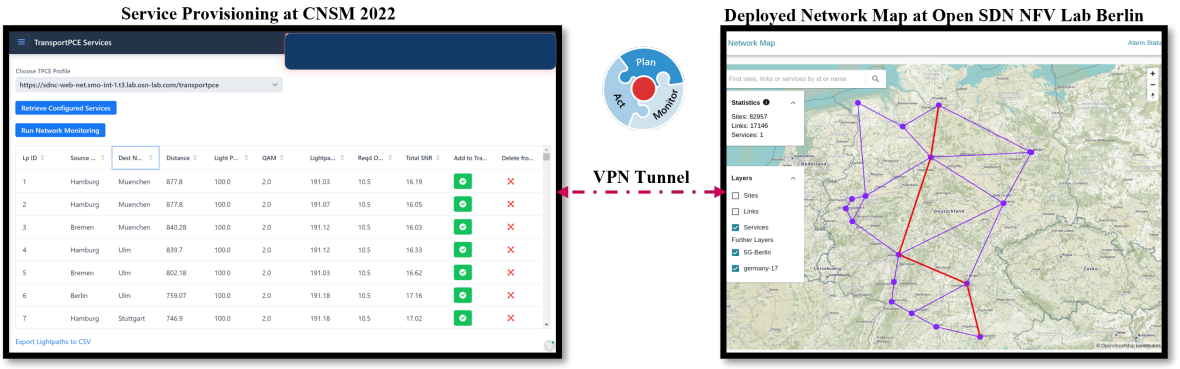


Figure 5.7: HeCSON WebApp deploys services remotely, which are visible in bold red on the network map of the simulated Nobel-Germany DOTN.

5.2.5 Summary and Conclusions

With ever-increasing traffic demands in OTNs, network operators aim to harmonize multi-vendor network elements. Using open-source data models like OpenROADM, coupled with an optical domain controller built for OTNs, end-to-end service planning, provisioning, monitoring, and restoration can be demonstrated. This solution can be deployed as a cloud-based offering, with the option to decouple various components. Inter-domain communication between the network orchestrator (*HeCSON*) and the optical domain controller (TransportPCE) is also demonstrated by using a VPN tunnel and TLS-enabled application layer secure REST APIs. This proof-of-concept opens development opportunities for network operators, to manage disaggregated optical networks. Future work involves the development of zero-touch provisioning and restoration functionality using streaming telemetry and deploying scalable network elements, which emulate real devices, thereby creating a digital twin of the entire optical network.

Chapter 6

Planning and Control of Quantum-Safe Optical Transport Networks

Optical networks, especially long-haul OTNs are traditionally considered secure from security attacks, since sensitive data propagated on optical channels is usually encrypted by asymmetric cryptography-based algorithms. This means that even if the optical fiber is "tapped", the stolen packets would contain encrypted information [127]. However, with growing interest and research in the field of quantum computing, solutions to computationally intensive problems can now be achieved in polynomial time. This also poses a threat to any information encrypted by cryptographic algorithms like RSA-4096 algorithm [128]. Therefore, quantum-safe solutions for securing highly sensitive data amongst both private and governmental organizations are needed.

OTN operators looking to future-proof their network deployments and assure ultimate network security, need to begin exploring different quantum-safe encryption methodologies and planning upgrades accordingly. Meanwhile, two quantum-safe solutions, namely QKD and PQC have emerged as commercially viable options to deploy a quantum-safe OTN. While PQC relies on lattice-based cryptography methods and can be deployed directly on the DWDM components by means of a software upgrade, QKD relies on deploying devices across the network to allow for information-theoretic secure exchange of symmetric keys between communicating entities [27].

This chapter answers the final research question **RQ4**, pertaining to the deployment and operation of quantum-secure solutions in existing long-haul OTNs. Since PQC does not require special considerations for physical layer network deployment, the first part of the chapter deals with the deployment strategies for QKDNs, given a pre-existing traditional OTN network topology. In the latter part of the chapter, a technology and vendor-agnostic

quantum-safe key management APIs is discussed, followed by a proof-of-concept for an SDN-enabled key.

6.1 Trusted Node Deployment Strategies in OTNs

QKD devices can be integrated into existing OTN infrastructure as discussed in Section 1.3. However, QKD devices have two limitations that need to be considered before deployment. Firstly, the QKD signal cannot pass through in-line amplifiers in the OTN network; and secondly, QKD devices have a reach limitation based on several physical and device-based parameters. Since OTN operators prefer to utilize pre-existing network infrastructure locations for cost reductions, each amplifier location in the network is treated as a candidate for QKD device placement. These locations need to have physical security and ensure that the QKD devices as well as key management servers are isolated and protected in every sense of the meaning. Such highly secure locations are a prerequisite to QKD transmission and are termed as **TRNs**. The first challenge that then must be tackled to deploy QKDNs is to minimize the number of TRNs and dark-fiber usage so as to ensure key-exchange capability across the network.

In order to evaluate various QKD network planning strategies, the effect of network deployments, number of trusted nodes, and types of QKD technologies on long-haul OTNs needs to be studied. To achieve this, a near-optimal span aggregation algorithm is introduced. This algorithm tries to minimize the number of TRNs on each link. Then, simulating the number of QKD demands in the network, a capacity planning study is undertaken. This study assumes a single QKD dark fiber pair between each TRN to compare different topological strategies. Finally, the effect of four different QKD technology types are compared and a multiple fiber capacity planning study is undertaken. The purpose of the multi-fiber capacity planning is to compare the number of deployed QKD fibers for each topological and technological combination.

6.1.1 State-of-the-Art Analysis

QKDN deployment over OTN was first studied over a specific network deployment, where a Minimum Spanning Tree (MST) and a single source shortest path-based heuristic was proposed [129]. However, no insights were provided on the applicability on different-sized reference networks. An “out-of-band” QKD deployment heuristic was presented in [130]. Since most of the networks operate a C-Band optical line system, deploying such an algorithm would also demand the installation of band splitters and combiners at each TRN location. Recently, field trials of QKD infrastructure based on untrusted repeater nodes have been

conducted. However, such a deployment is suitable only for limited-range metropolitan networks [131].

For planning QKDNs with existing technologies, two mixed-integer linear programming-based solutions are introduced to minimize the number of parallel QKD fibers in the network [132]. However, this work assumes that the trusted nodes are pre-assigned and known to the planning algorithm. Further, the resulting solution space is vast which leads to large computation times.

Finally, Li *et. al.* introduce a mathematical model for topology evaluation of QKDNs [133]. Using a flow-based model, two long-haul networks are simulated and evaluated in terms of the achievable secure key rates. Despite providing an initial algorithm for adding intermediate nodes to the QKD chain, the authors treat the study of relationship between QKD device reach distance, topology, and other link parameters as part of future work.

Overcoming these challenges, this section highlights the practical challenges of deploying a TRN-based QKDN over existing OTN infrastructure and examines strategies used to reduce the number of TRNs in the network. Since every method to reduce TRNs would lead to lower secure key rates, the compromise of lowering the number of TRNs in the network with the decrease in the achievable secure key rates in the network is studied.

6.1.2 QKD Deployment Strategies

To plan a QKDN, the fiber topology of the existing OTN network needs to be available. Estimates on fiber span lengths can be made based on open-source fiber deployment data, as well as span length optimization experiments [134]. Using such data, a fiber-span length model with a mean of 80 km (optimized for long-haul optical transmission) and a standard deviation of 10 km is created. The standard deviation caters to irregularities in span lengths that may occur due to geographical or administrative constraints. To generate fiber span information for each link in the networks under study, namely, Nobel-Germany (Fig. 6.1a) and Nobel-EU (Fig. 6.1b) [30], span lengths are drawn from the modeled distribution.

From a network capacity planning perspective, the choice of network topology, the placement of trusted nodes, as well as choice of QKD devices need to be explored to find a solution that is suitable for each network. Therefore, two deployment strategies are broadly defined and evaluated; namely *topological* and *technological* strategies.

6.1.2.1 Topological Strategies

The existing OTN infrastructure consists of traffic-generating nodes called *PoPs* and amplifier nodes called ILA. PoPs and ILAs are connected using several bundles of optical fiber cables.

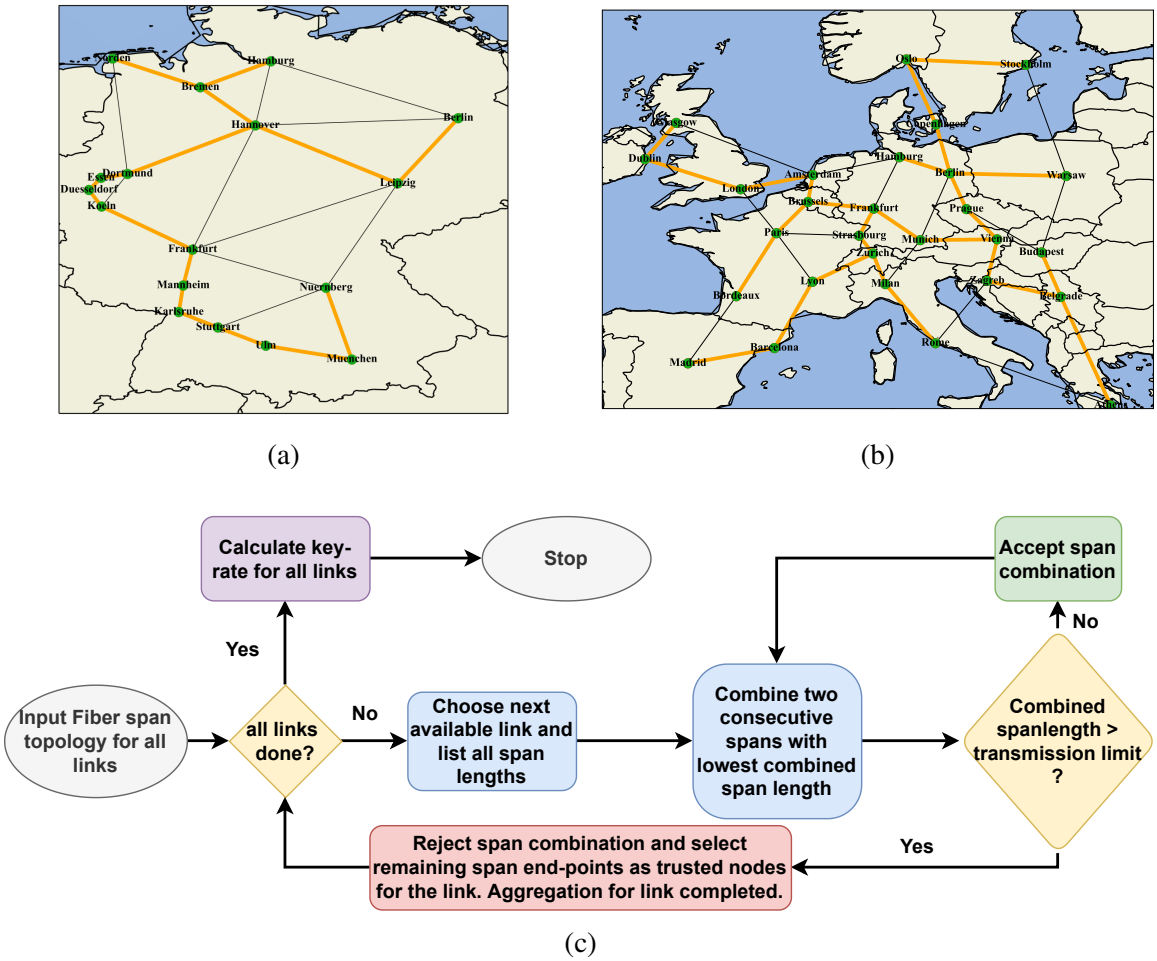


Figure 6.1: (a) and (b) Nobel-Germany and Nobel-EU topology with MST (yellow highlighted) respectively, (c) Span aggregation algorithm flowchart.

The unused optical fibers between any two ILAs or a PoP and an ILA are called dark fibers. Further, a consecutive collection of optical fibers connecting any two PoPs is defined as a link.

Since Secure Application Entitys (SAEs), which generate key requests are located only at PoP locations, there are two ways of routing these demands between any two PoPs. The first is to use any available link in the network to find the shortest path. This method is addressed as ***baseline*** in this work. The second method is to use only links on the network's minimum spanning tree based on link lengths (***MST_dist***). As discussed in [129], MST_dist reduces the number of links on which trusted nodes need to be deployed thereby bringing down the network cost. The disadvantage, however, is the lack of protection paths and reduced SKR. Figs. 6.1a and 6.1b show the MST links for Nobel-Germany and Nobel-EU networks respectively.

Since the assumption that every ILA location is co-located with a trusted node potentially increases the expected operation and installation costs for operators, a span aggregation (SA)

algorithm is introduced, as described in Fig 6.1c. Further, two strategies, namely ***baseline_SA*** and ***MST_SA*** are created, which use the span aggregation algorithm on the baseline and MST topologies respectively.

As described in Fig. 6.1c, the span aggregation algorithm can be applied to all links in the given network topology. First, all the fiber spans of the link are listed. Then, using a binary-search method, two consecutive spans with a minimum combined span length are chosen. The combined span length is then verified to be below the transmission limit of the QKD device [29]. If that is the case, the span aggregation is accepted and the intermediate ILA is bypassed. This process continues, till a combination of span length results in exceeding the transmission limit. At this stopping condition, the aggregated span end-points are assigned as TRNs and the link is considered aggregate.

Deployment Strategy (Topological)	**Topology type**	**Span aggregation**
baseline	Full network	None
baseline_SA	Full network	ref. Fig. 6.1c
baseline_full	Full network	Full-depth search
MST_dist	Minimum Spanning Tree	None
MST_dist_SA	Minimum Spanning Tree	ref. Fig. 6.1c
MST_dist_full	Minimum Spanning Tree	Full-depth search

Table 6.1: Overview of topological deployment strategies.

Finally, since the span aggregation algorithm is suboptimal, a full-depth search method to minimize the number of TRNs on each link is also implemented. The full-depth search gives an optimal solution for sapn aggregation by combining consecutive span lengths till the stopping condition of QKD device transmission limit is reached. The deployment strategies ***baseline_full*** and ***MST_full*** implement the full-depth search on the baseline and MST topologies respectively. Table 6.1 summarizes the discussed strategies. Hence six different topology-based deployment strategies are evaluated on two networks under study and compared in Section 6.1.4.

6.1.2.2 Technological Strategies

In recent years, long-haul applications for QKD transmission have led to the development of commercially available QKD devices which claim transmission distances of up to 150 km, while ensuring SKR of at least 1 kbps [29]. It is assumed that the ability of a QKD device to transmit at longer distances is directly proportional to its cost. Therefore, four different types of QKD technologies can be envisioned as follows:

1. LC1 - The first low-cost device assumes QKD BB84 Decoy-state protocol with fiber attenuation of 0.2 dB/km, dark-count rate of 10000 s^{-1}, APD quantum efficiency of 20% and a repetition rate of 1.25 GHz.

2. LC2 - The second low-cost device assumes QKD BB84 Decoy-state protocol with fiber attenuation of 0.22 dB/km, dark-count rate of 10000 s^{-1}, APD quantum efficiency of 20% and a repetition rate of 1 GHz.

3. HC1 - The first high-cost device assumes QKD BB84 Ideal protocol with fiber attenuation of 0.2 dB/km, dark-count rate of 10000 s^{-1}, APD quantum efficiency of 20% and a repetition rate of 1.25 GHz.

4. HC2 - The second high-cost device assumes QKD BB84 Ideal protocol with fiber attenuation of 0.22 dB/km, dark-count rate of 10000 s^{-1}, APD quantum efficiency of 20% and a repetition rate of 1 GHz.

The design choice of QKD protocols, the range of fiber attenuation constant, as well as QKD device parameters are obtained from product data sheets of QKD devices [29, 52]. For HC1 and HC2, using BB84 Ideal protocol [53] is justified as a stop-gap for better algorithms and advancements in photonic transceivers. As seen in Fig. 6.2, each of the QKD devices has

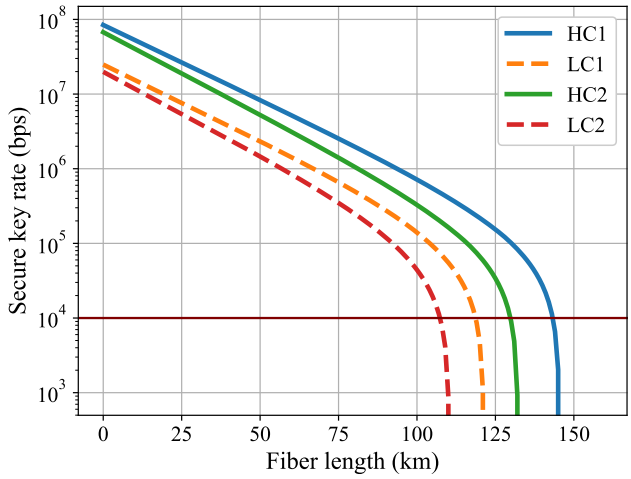


Figure 6.2: SKR versus fiber length for four different QKD technologies [55].

an SKR value associated with the fiber length. Therefore, a cut-off length is also defined for each of the four technologies. This cut-off length is used as an input to the span aggregation algorithm. To derive the cut-off length, the lower-bound on the SKR of each fiber span is fixed to 10 kbps. This implies that none of the spans in the network deployed with LC1, LC2,

HC1, and HC2 can be greater than 120, 105, 140, and 130 km respectively. The impact of the four different QKD devices is analyzed in detail in Section 6.1.4.

6.1.3 QKD Network Capacity Planning

In order to evaluate the topological and technological strategies, a methodology for QKDN capacity planning is devised. As previously explained in Section 1.3, SAEs request keys from KMEs, which in turn refresh their key store by symmetrically loading keys from each QKD device pair. The generation of these keys is highly dependent on the achievable SKR of each QKD fiber pair along a chosen path in the network. Therefore, the addition of SAE requests in the network is simulated and each incoming request is routed onto the QKD fiber pairs. The objective of the capacity planning is to accommodate as many SAE requests as possible in the network, for a given topological and technological (LC1, LC2, HC1, HC2) strategy.

The capacity planning considering only topological strategies assumes a single fiber pair between any two TRNs whereas the capacity planning considering both topological and technological strategies allows for the addition of new QKD fiber pairs. For the multiple fiber capacity planning, in case any QKD fiber pair serving pre-existing SAEs reaches its maximum capacity, a new QKD fiber pair is added between the TRNs.

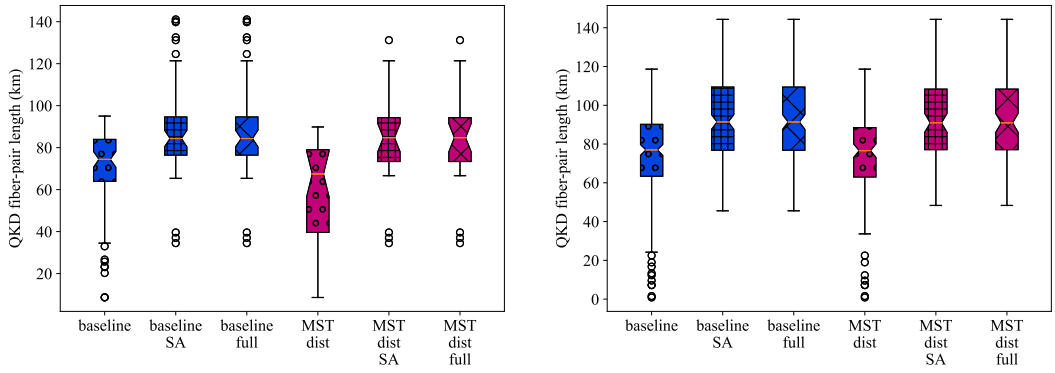
To simulate the effects of adding SAE requests to the six topological deployment strategies, a QKD demand simulator is implemented, which first finds up to $k = 5$ shortest paths for the baseline topologies and the first shortest path for the MST_dist topologies for all end-to-end PoPs in the network. Then, the SAE requests are sorted from the longest to shortest paths. To simulate the requested SKR for each SAE, a Gaussian distribution with a mean of 100 kbps is used. Similar to [130], it is assumed that SAE requests have an infinite holding time. Each simulation run, however, is restricted to 1000 SAE requests.

Each incoming SAE request is routed on the path with the highest capacity out of the k-shortest paths thereby avoiding greedy filling of the first shortest path links. In case of the single-fiber planning, the SAE request is dropped if it cannot be assigned to any of the paths. For the multi-fiber approach, after each SAE request is added, if any QKD link has an available capacity of less than 5% of its maximum capacity, a parallel QKD fiber-pair is deployed on the bottleneck span. Therefore, the multi-fiber approach minimizes the probability of dropped SAE requests while increasing the number of QKD fiber-pairs and devices.

To improve confidence in our results, 100 random simulations are run for each comparison type. The evaluations have been performed on a machine equipped with 11th Gen Intel® Core™ i7-1185G7 @ 3.00GHz, 32 GB of RAM, running Windows 10. The time taken for each simulation run is in the order of milliseconds.

6.1.4 Results and Discussion

In this section 6 topological strategies are compared in terms of the cumulative SKR for Nobel-Germany and Nobel-EU networks. Here, the effect of span aggregation algorithms is also shown for each of the networks across different topological strategies. Then, assuming that the proposed span aggregation algorithm is used on the *baseline* and *MST_dist* topologies, a combination of 8 topological and technological strategies are compared in terms of the number of QKD fiber pairs and the number of QKD devices in the network.



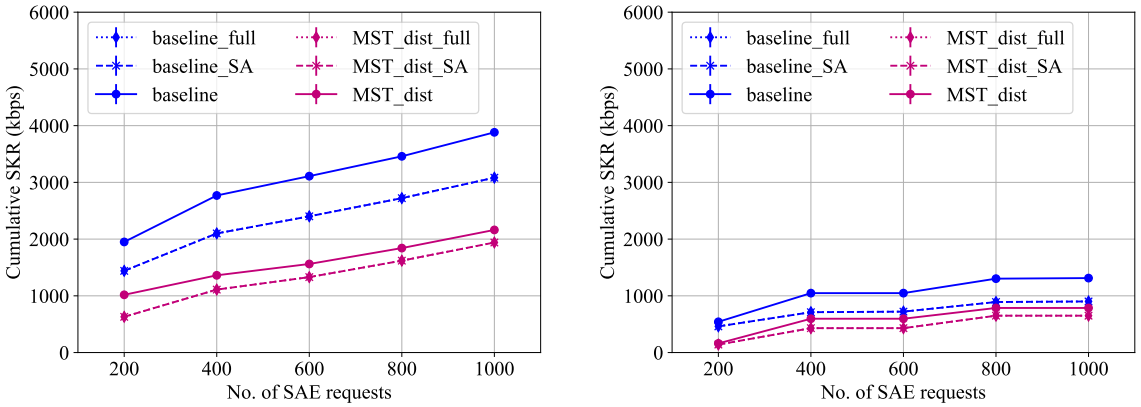
(a) Nobel-Germany (b) Nobel-EU

Figure 6.3: Span length distribution for networks under to show the effect of span aggregation, fixing the technology strategy to HC1.

6.1.4.1 Topological Comparisons

Figs. 6.3a and 6.3b show the effect of span aggregation on each of the topological strategies. For Nobel-Germany, span aggregation leads to a mean increase of 15% and 18% in the length of fiber-pairs between TRNs as compared to the *baseline* and *MST_dist* strategy respectively. For Nobel-EU a similar increase in the fiber-pair lengths is observed.

From the results, it is observed that although span aggregation techniques place lower SKR at higher SAE requests, the cumulative SKR is comparable to the non-aggregated topologies when SAE requests are lesser in the network. Therefore, a strategy of first deploying a span aggregation algorithm based solution and then adding additional TRNs can be planned, for example first deploying *baseline_SA* and then progressively moving towards *baseline* as SAE requests in the network increase. Moreover, comparing the span aggregation algorithm with the full-depth search, the results for both across all topologies and networks are similar. This is since the span lengths have a mean of 80 km, there are very few consecutive span lengths in the network which are aggregated, leading to negligible advantage of using a full-depth search for span aggregation.



(a) Nobel-Germany (b) Nobel-EU

Figure 6.4: Cumulative SKR vs. number of SAE requests for single QKD fiber-pair per link deployment.

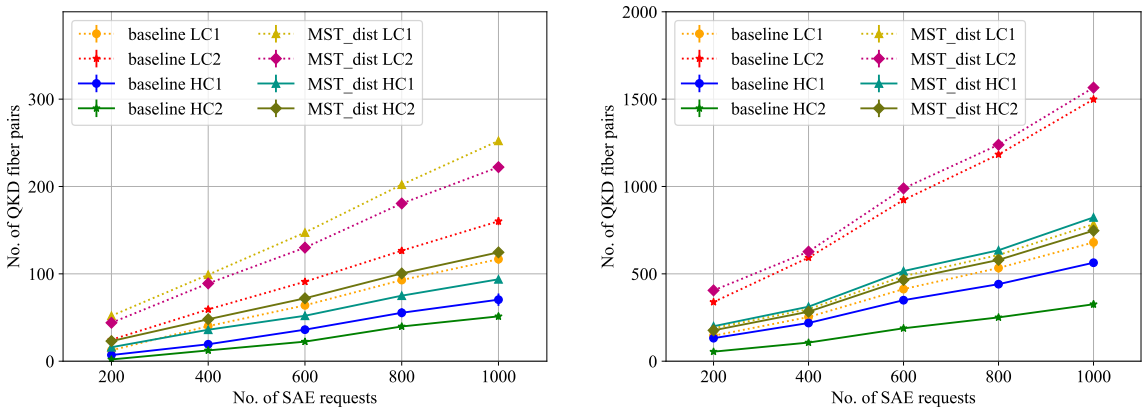
Figs. 6.4a and 6.4b show the cumulative SKR carried by the Nobel-Germany and Nobel-EU network respectively, versus the number of SAE requests added into the network simulator. In this study, HC1 (transmission limit 140 km) is the choice of QKD device for all the simulations and only one QKD fiber-pair is available between every TRN location.

For both the networks under study, *baseline* strategy provides an upper-bound on the cumulative SKR. However, it is also the most expensive option since all the ILAs in the network are converted into trusted nodes. The *baseline_SA* and *baseline_full* as well as *MST_dist_SA* and *MST_dist_full* strategies have the same SKR because they have the same span lengths. As compared to *baseline*, *MST_dist* places almost 50% lesser SKR into the network. Since all the SAE requests in an MST topology can only be routed on their first shortest path, the link capacity saturates early, resulting in lower cumulative SKR. It is interesting to note that *baseline_SA* performs only slightly better than *MST_dist* in the case of Nobel-EU network. This is because span aggregation, while helpful in reducing the overall number of devices, can also reduce the SKR, despite having path diversity. For a given topology with span aggregation place on an average 20% lesser cumulative SKR at the maximum number of SAE requests.

6.1.4.2 Technological Comparisons

After evaluating various topological strategies, the effect of four different technologies is evaluated (as introduced in Sec. 6.1.2.2) on the networks under study. Since from Fig. 6.4 it is clear that the proposed span aggregation algorithm works identically to the full depth search on the given networks, only the topologies with span aggregation are considered in this study, namely, *baseline_SA* and *MST_dist_SA*. For brevity, the suffix "*SA*" is ignored from hereon.

In order to evaluate the benefits of LC1, LC2, HC1, and HC2, it is ensured that minimum SAE requests are blocked by using the multiple fiber scenario (described in Sec. 6.1.2.2). For all the 8 scenarios evaluated, the cumulative SKR, as well as the number of SAE requests placed are similar. Therefore, the number of fiber pairs placed are compared and inferences for deployment planning are drawn.



(a) Nobel-Germany

(b) Nobel-EU

Figure 6.5: QKD fiber pairs vs. number of SAE requests.

As seen in Figs. 6.5a and 6.5b, in terms of number of QKD fiber pairs placed, *baseline HC2* places the least number of fiber pairs. We also note that the minimum spanning tree-based topologies, namely *MST_dist LC1*, *MST_dist LC2*, *MST_dist HC1* and *MST_dist HC2*, consistently place more fiber pairs than their *baseline* counterparts using the same technology type in the Nobel-EU network.

From the technological analysis, for Nobel-Germany network, *baseline HC2* emerges as a clear choice for deployment. However, if the number of SAE requests stay relatively low, operators can also aim for an *MST_dist HC2* deployment thereby reducing the overall number of TRNs to be operated and maintained in the network.

6.1.5 Summary and Conclusions

TRN deployment strategies aim to enable QKDN deployment in long-haul OTN. These strategies combine topological changes and a novel span aggregation algorithm to reduce the total number of TRNs in the network. The span aggregation algorithm allows for up to 20% higher QKD fiber pairs when the technology strategy is pre-selected. The proposed span aggregation algorithm, although sub-optimal, shows similar results as compared to an optimal full-depth search algorithm. Of the eight topological and technological strategies, a full network deployment of QKD, while using the span aggregation algorithm, and deploying the

HC2 flavour of QKD devices (*baseline HC2* in Figs. 6.3 and 6.5) results in a saving of at least 3 times the number of fiber-pairs deployed as compared to the worst performing topological and technological combination (*MST_dist LC2*). Although QKDN deployment planning is a complex task with several degrees of freedom, comparative studies provide strategic insights to operators planning to deploy similar networks.

Future work will look into a mixed deployment and an upgrade strategy amongst the different topological and technological combinations. As the results are highly dependent on the network as well as the cost of devices and fiber pairs, a detailed cost model for a techno-economic analysis is also under development.

6.2 Software-defined Key Management for Quantum-Safe Networks

As traditional DWDM networking components have been integrated with SDN-based control, it is natural to expect similar separation of management and control planes for quantum-safe key generating and key management devices. Here two aspects of enabling SDN-based solutions for quantum-safe networks are discussed. The first aspect deals with the implementation and development of a technology-independent cryptographic module with a key delivery interface in quantum-safe OTNs. Since the key delivery interface should work with any quantum-safe solution, we discuss the implementation not only for QKD interfaces but also for PQC interfaces

The second aspect is the demonstration of a key management methodology for QKD devices using commercial devices, which can interact with different QKD devices using an ETSI standard interface south-bound and vendor-specific REST API north-bound to interact with an optical domain controller.

6.2.1 State-of-the-Art Analysis

Early attempts to integrate SDN methodologies with QKD and PQC devices have been attempted as part of field trials and demonstrations. Ou *et. al.* demonstrated an OpenDayLight-based controller integrated with QKD devices in a field trial of a QKDN [135]. Aguado *et. al.* discussed the architectural implications and functionalities of an SDN controller in QKDN and showcased this on a three-node QKDN deployed in a metro region. Both works, however, use non-standard APIs for key exchange, leading to limited solution reproducibility.

There have been several works explaining the key delivery interface with PQC as a part of the Open Quantum Safe project [136–138]. The cryptographic libraries developed by this

project have now been standardized and are used in several key-delivery algorithms. However, at the time of the publication of this work, and to the best knowledge of the authors, there were no works that provided combined key delivery from QKD devices and PQC encryptors into a KME. Further, the management of the key delivery path using an SDN contoller was previously limited only to QKD devices. However, with this work, it is extended into a technology and vendor agnostic key delivery and management solution.

6.2.2 Key exchange APIs for PQC and QKD networks

This work demonstrates a framework of quantum-secure solutions deployable on existing optical networks in a flexible and scalable manner. In the presented framework, both QKD and PQC are handled by a **standard key interface**. Encryption keys are accessed from a key server of either QKD or PQC, depending on the user requirement, to data encryptors in a unified key interface.

Similar to a standalone QKD system, a PQC key exchange protocol is executed on a dedicated cryptographic module (known as $CryptoM$). $CryptoM$ can host multiple key exchange processes and supply session keys to data encryptors upon their requests. Each session key is identified by a keyID and strictly tied to a specific data encryptor.

Furthermore, in order to minimize the latency of key delivery, key exchange protocols are periodically performed and encryption keys are generated before requested. In this scenario, secret keys need to be stored in a server, which has actually a potential risk of a single point of failure. To eliminate such risk, a $(t,\ n)$-threshold key management protocol is implemented.

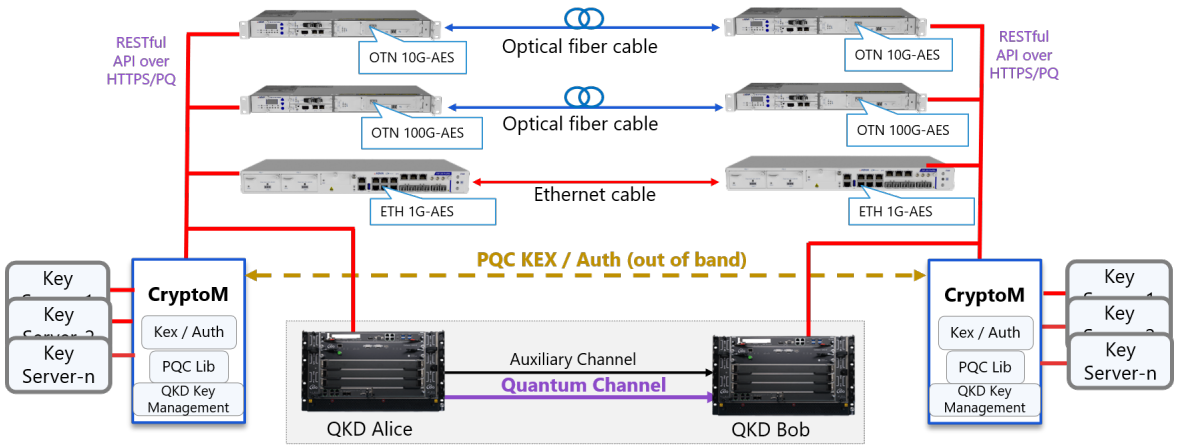


Figure 6.6: System architecture of the demo system (© IEEE/OSA 2021).

The three building blocks needed to build a KME are user data encryption, authenticated key supply, and threshold key management, each of which targets 128-bit quantum security.

A block diagram of the demo system shown in Fig. 6.6 includes the three building blocks, which are described as follows:

1. Data encryption: User data are encrypted using a symmetric-key crypto algorithm such as AES-256-GCM, which suffices the targeted security levels.

2. Authenticated Key Supply: An encryption key is derived by executing either a QKD protocol such as BB84 or a PQC key exchange protocol such as IKE. The derived key is delivered to encryptors via REST APIs over Hypertext Transfer Protocol (Secure) Post-Quantum (HTTPS/PQ).

3. Threshold Key Management: Encryption keys are not stored in $CryptoM$ (ref. Fig. 6.6), rather they are converted into n shares and distributed to multiple key servers. When requested, $CryptoM$ collects n shares from the key serves and reconstructs the original key by a threshold key management protocol. An encryption key is correctly reconstructed as long as at least any $t + 1$ out of n shares are correct.

6.2.3 Post-Quantum Cryptography and Key Delivery Interface

Through the course of this dissertation, the focus has been on QKD devices, their planning and management. However, a key delivery interface and a KME needs to accept keys from both QKD as well as PQC-based encryptors. Therefore, it is pertinent to discuss the architecture and techniques of key delivery for PQC interfaces also.

SDO	PQ KEM	PQ Signature
NIST	Classic McEliece CRYSTALS-KYBER NTRU SABER	DILITHIUM FALCON Rainbow
IETF	-	XMSS LMS

Table 6.2: PQC primitives; the third round finalists from NIST [139] and hash-based signatures from IETF [140].

Although the standardization process is currently ongoing, the 3rd round finalists of NIST PQC project would be the best candidates for PQC key exchange [139]. They are listed in Table 6.2. In addition, hash-based signatures should be accounted for, since they have been already standardized in IETF [141, 142] and are supported by NIST [140]. Note that KEM stands for Key Encapsulation Mechanism defined as a technique by which a data encryption key is derived. Signature schemes are typically used for entity authentication. All PQC primitives (or algorithms) shown in Table 6.2 are supported in the proposed implementation.

HTTPS is a widely used secure communication protocol based on TLS and it is used for the QKD key delivery in the ETSI standard [27]. HTTPS/PQ is a hybrid version of the HTTPS protocol based on a classical cryptosystem such as ECDHE-RSA and a PQC system such as NTRU-Dilithium in Table 6.2. The REST APIs provided by ETSI for QKD devices are extended to work with PQC devices, using an HTTPS/PQ protocol.

API Name	Description
GET Key	Request a key from the key supplier
GET Key with keyID	Request a key that matches the given keyID
GET Status	Request the status of keys

Table 6.3: REST APIs for key delivery from consumer to supplier (© IEEE/OSA 2021).

In this work, a set of REST APIs listed in Table 6.3 are implemented to access the status of keys or a key that is staged in QKD or $CryptoM$. The authenticated key exchange protocol between an encryptor i at NCU_A and encryptor j at NCU_B is shown in Fig. 6.7 and briefly explained as follows[1]:

1. NCU_A sends a GET Key REST API to $CryptoM_a$.
2. $CryptoM_a$ and NCU_A mutually authenticate.
3. $CryptoM_a$ sends a secret key with its keyID to NCU_A over HTTPS/PQ.
4. NCU_A transfers the keyID to NCU_B.
5. NCU_B sends a GET Key with keyID API in the request body to $CryptoM_b$.
6. $CryptoM_b$ and NCU_B mutually authenticate.
7. $CryptoM_b$ sends a secret key to NCU_B over HTTPS/PQ that the keyID is matched.

Once an authenticated key exchange protocol is completed, a secret key is shared in both $CryptoM$ modules. Then, this secret key is converted into n shares using Shamir's secret sharing scheme [143] and distributed to n key servers. When a secret key is requested, $CryptoM$ aggregates at least $t + 1$ shares from the key servers, reconstructs the secret key and delivers it to the user. After delivery, the secret key is completely deleted from the key servers and is never used again.

To instantiate this scheme, a Shamir secret sharing scheme with $(t, n) = (2, 5)$ is implemented. A degree-2 random polynomial $q(x) = a_2x^2 + a_1x + a_0$ is generated and $a_0 = \kappa$ is set. To share the secret, five points on the curve are randomly chosen and distributed to five

[1]NCU stands for Network Control Unit

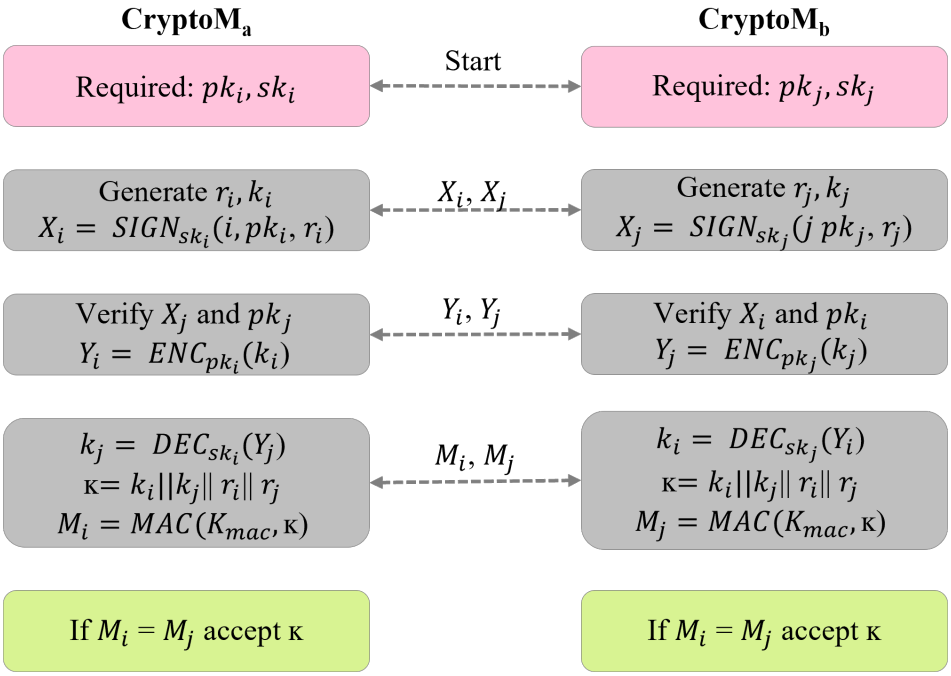


Figure 6.7: Authenticated Key Exchange protocol.

key servers, respectively. Hence, it is possible to reconstruct the $q(x)$ if any three points are correctly collected.

The described implementation can be extended for $n > 5$. However, the complexity of the reconstruction process is exponentially increased accordingly since every n combination of shares should be verified before a correct key is derived.

6.2.4 SDN-enabled Key Management Entity

To support the interoperability of QKD devices from different vendors, a KME on a trusted node is functionally decoupled from the QKDN. The KME accesses QKD keys individually via the ETSI standard key delivery interface. Each KME can collect keys from intermediate QKD devices, encrypt them by using a one-time pad, and deliver them to the destination node [55]. A key generated at the source node can be relayed through a chain of QKD links which are dynamically allocated by an SDN controller. The key is restored at the destination node by XORing all the ciphertexts received from the intermediate nodes [27]. Trusted nodes are mutually authenticated by a quantum-secure authentication method which is immune to retrospective decryption by future quantum computers.

The proposed key management system is implemented on commercial products which have multiple network interfaces. An example of the key management solution with the XOR-based key relay method is shown in Fig. 6.8.

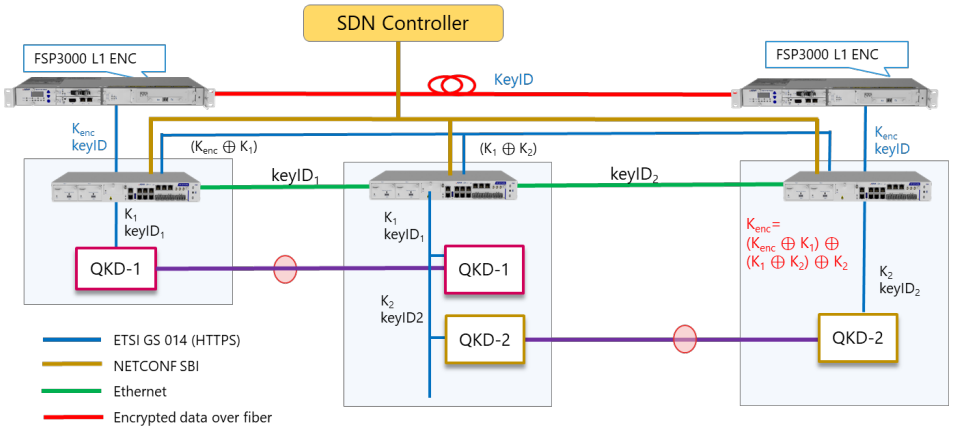


Figure 6.8: Configuration of Key Management using two QKD links (© IEEE/OSA 2021).

The proposed KME can be implemented at each TRN (light blue boxes in Fig. 6.8) in the QKDN. This solution has the potential to integrate an unlimited number of TRNs for a meshed QKDN. The novelty is that the KME implements standard interfaces defined by ETSI, which allow different QKD device vendors to be connected to the KME. This enables disaggregation at the QKD device layer, allowing OTN operators to choose different QKD devices at different TRNs. For each source-destination QKD demand, a key relay route can be pre-calculated using the techniques discussed in Section 6.1.3 and stored in the SDN controller. As shown in Fig. 6.8, the SDN controller needs to be connected to all the KME devices in the QKDN.

This demonstration is composed of commercial products available in the market. The QKD devices are supplied by IDQuantique [29], while the KME is deployed on ADVA FSP150™ [18], as seen in the block diagram in Fig. 6.9. To enable XOR-ing, the keyID of the QKD key stored in a given TRN is sent to the next node via an Ethernet port (N1 in Fig. 6.9). The key relay route can be dynamically switched by an SDN controller, depending on the available symmetric keys in QKDN. In case symmetric keys are not available, this system can also be used to relay traditional keys, as long as the keys are inserted into the KME's database (Key DB (SQL) in Fig. 6.9). The KME then interconnects all QKD links between the source and destination, while XOR-ing the keys along the way. Hence, SAEs also have the option to insert their own key and let it relay over the QKDN.

Using QKD clients in Fig. 6.9, a node-to-node communication is performed via the ETSI GS 014 southbound interface. The one time pad encrypted key is delivered to the destination node or the respective SAE, using an HTTPS-based server. TRNs are mutually authenticated by verifying their Public Key Infrastructure (PKI) certificates before the key is delivered. The PKI certificate is produced using the post-quantum PKI.

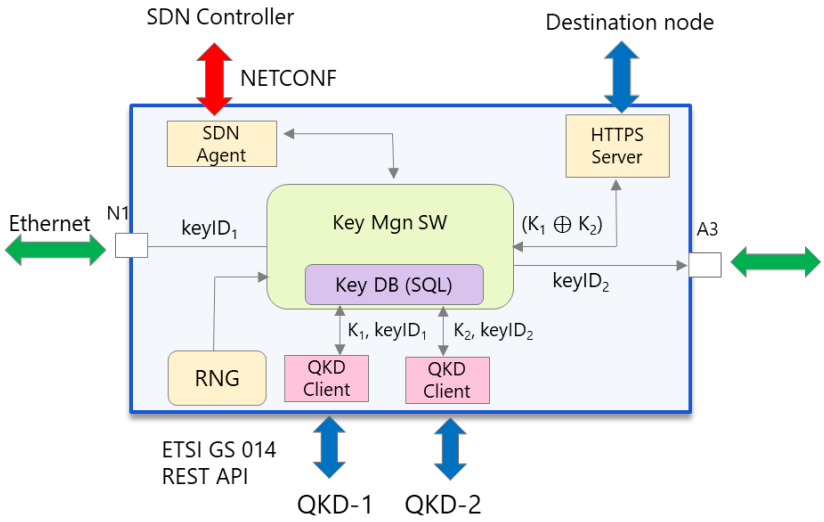


Figure 6.9: Block diagram of the implemented key management agent (© IEEE/OSA 2021).

A built-in key database (Key DB (SQL) in Fig. 6.9) provides the capability of buffering the one time pad encrypted QKD keys to speed up the key relay and reduce the latency. Each individual software component is implemented as an independent linux-based service.

In a meshed QKDN, a centralized SDN controller oversees dynamically provisioning and rerouting key relay routes, thus fulfilling the QKD service or network robustness requirements. The controller can poll key relay information from trusted nodes using NETCONF as a south-bound interface in conjunction with proprietary YANG models available from each of the commercial products.

6.2.5 Summary and Conclusions

Even though quantum computers are still under research, deployment of new solutions against quantum threats is already being deployed to secure communication infrastructure. However, it is quite challenging to implement such quantum-safe solutions on existing devices, in particular, on optical networking systems which are based on a resource-limited embedded systems platform. In this section, a framework of flexible and scalable quantum-resistant encryption combined with threshold key management is presented. Further, the scalability of this solution is increased by deploying an simple SDN-controller which provides a standardized interface to the deployed KMEs. Future work involves the use of machine learning for dynamic key threshold management and ensuring that symmetric keys are available as the number of SAE requests in the network increase.

Chapter 7

Conclusion and Outlook

In this dissertation, four research questions related to the planning, control, and security of Optical Transport Networks (OTN) have been answered. The trend of deploying flexible, scalable, and disaggregated solutions is envisioned to improve interoperability and drive down the total cost of ownership. Such flexible and quantum-safe OTNs also form a crucial backbone for long-distance low-latency communication between geographically distant regions and are central to the 6G networking paradigm.

Therefore, the contributions of this work will help interested stakeholders to plan, deploy, operate, and protect their digital infrastructure, thereby offering robust communication channels for various applications. In this chapter, each of the contributory chapters of this dissertation is briefly summarized and individual conclusions are drawn. In the end, avenues for future work and extensions to the contributions are discussed.

7.1 Summary and Discussion

Here we summarize each chapter of this dissertation that contributes to answering the research questions.

7.1.1 Optical Transport Network Planning Algorithms (RQ1)

In Chapter 3, a network planning heuristic for configuration selection in flex-grid enabled OTNs is introduced (*HeCSON*). This work highlights the motivation and advantages of using flex-grid enabled Bandwidth Variable Transceiver (BVT)s in an OTN, thereby increasing the network throughput and spectral density for present-day networks.

Then, we extend this work by proposing a multi-period planning combinatorial optimization based algorithm (*MP-HeCSON*), which optimizes the number of BVTs needed by each

demand in each planning period. *MP-HeCSON* also has a traffic reactive ability and changes the number of BVTs deployed in the network according to the demand network.

Finally, we explore a regenerator-based BVT allocation workflow as an extension to *MP-HeCSON*, which allows for a greater increase in overall network throughput and spectral utilization.

Each of these algorithms is tested on a single fiber-pair C-band OTN simulations of different kinds of networks under study. It can be concluded that although Routing, Configuration, and Spectrum Allocation (RCSA) is an NP-hard problem, using meta-heuristics in a divide-and-conquer fashion can provide meaningful insights into future network deployments.

7.1.2 Multi-band Network Planning Studies (RQ2)

Chapter 4 begins by comparing the C-band-based network studies done previously with multi-band network scenarios, i.e., considering that L-band and S-band on each optical fiber in the network can be activated. Although multi-band transmission needs newer equipment at every Points of Presence (PoP) and In-line Amplifier (ILA), it can be concluded that the savings on dark fiber can be made and are especially beneficial to OTN operators who lease dark fiber.

Another aspect of multi-band planning depends on the performance of different kinds of terminal devices and the initial choice of using them in the network. Simulations show that due to their robustness in the optical domain, BVTs can carry 20% more traffic than transparent Internet Protocol over Wavelength Division Multiplexing (IPoWDM)-based pluggables. Although these pluggables are lower in cost, they require upgrades from grey-optics-based IP routers to coherent IP routers, leading to an increase in overall Capital Expenditure (CAPEX).

To understand how the migration to a new band or a new dark fiber affects the cost, we devise an Integer Linear Programming (ILP), as well as a Reinforcement Learning (RL) based heuristic to conduct a network migration study. It is observed that the RL solution performs as well as the ILP solution, with minimal changes in the overall traffic placed, but with lower execution times, showing the benefits of using RL for quicker network analysis.

With this, we conclude our discussions on network planning and deployment for OTNs. While multi-band deployments have already begun, there is potential in further exploring the different bands available in the optical fiber for optical channel propagation, thereby delaying the laying of additional fibers in the ground, at least in the near future.

7.1.3 Monitoring & Control of Disaggregated Optical Networks (RQ3)

In Chapter 5, we discuss several statistical and Machine Learning (ML) based algorithms to detect soft failures in BVTs. These soft failures were generated on BVTs deployed on a live production OTN and the data was collected over 45 days. While the ML models showed great accuracy, there is a risk of overfitting and false positives in such an approach. Nevertheless, the potential of using ML models to reduce complexity and increase network monitoring potential is greatly appreciated.

Further, since network disaggregation gains traction, it is important to test disaggregated control plane solutions on a network simulation and optimize several workflows to minimize traffic disruptions in open OTNs. A proof-of-concept based on OpenROADM compliant simulated Reconfigurable Optical Add Drop Multiplexer (ROADM)s and Optical Transponder (XPDR)s, connected to an optical domain controller TransportPCE via NETCONF is presented. TransportPCE exposes Representational State Transfer (REST) Application Programming Interface (API), which are consumed by a higher layer network planner and orchestrator to plan, add, remove, and reconfigure optical services in the network, thereby leading to effective and open-source OTN management.

7.1.4 Planning & Control of Quantum-Safe Optical Transport Networks (RQ4)

Chapter 6 deals with the deployment and control of quantum-safe keys into the traditional OTN. Two quantum-safe technologies, namely Quantum Key Distribution (QKD) and Post-Quantum Cryptography (PQC) are explored.

From a deployment perspective, several Trusted Node (TRN) deployment strategies for long-haul Quantum Key Distribution Network (QKDN) are explored. The proposed span aggregation algorithm allows for cost savings by skipping many candidate TRN locations. From the various deployments studied, it can be concluded that QKDN operators will benefit from a staggered deployment, that is, first starting with QKD deployments on a Minimum Spanning Tree (MST) and then gradually activating QKD on other links to accommodate higher Secure Application Entity (SAE) requests.

Even though QKD and PQC solutions have been extensively deployed in field trials and demonstrators, commercial deployment is subject to standardization. To ensure smooth integration of quantum-safe key delivery interfaces with minimal implementation on the OTN operator's side, a technology, and vendor-agnostic key management framework is envisioned. This key management framework consists of a Key Management Entity (KME) which interacts with QKD devices and PQC-based encryptor modules ($CryptoM$), to exchange quantum-safe

keys in a long-haul OTN network. Such a scalable solution uses standardized REST APIs as defined by ETSI and a post-quantum enabled transport layer security (HTTPS/PQ), assuring solution compliance and device interoperability.

7.2 Future Work

Algorithms to plan OTN deployments have developed from very simple throughput maximization based RSAs to multi-band multi-fiber multi-period RCSAs. At the same time, with the push towards standardized control plane characteristics and open-source-based data models, OTN device manufacturers now need to offer complex solutions on the physical layer and network layer, while ensuring that the cost expectations of the OTN operators are not breached.

For network planning algorithms, ML models can be further explored to increase accuracy and reduce execution times for several sub-tasks within the RCSA framework, for example, physical path calculation or spectrum allocation. Transparent IPoWDM solutions are getting increasingly robust and are significantly cheaper than BVTs. Extensive studies with both solutions must be conducted before choosing either technology as a candidate for a long-term OTN deployment.

With the applicability and success of a digital twin in several other industries, it also makes sense to apply this concept to OTNs. Creating a digital twin of an OTN has several advantages, including the possibility of testing new scenarios in a "sandbox", before deploying on the live production network. The distinguishing factor between a normal simulation and a digital twin is the ability of the twin to mimic current network behavior since each "twin" is continuously updated with data from real devices using streaming telemetry.

Finally, as quantum-safe devices like QKD and PQC reach the deployment phase, research must look forward to building solutions for next-generation quantum-safe devices like out-of-band QKD and entanglement-based QKDs. Such planning and control algorithms need to be also secured from quantum computer attacks, and therefore the paradigm of "quantum-safe SDN" must be explored.

Appendix A

Gaussian Noise Models for Network Planning

Here we provide an overview of each of the GN models used in this dissertation. GN models are a class of perturbative models that have been widely used to accurately predict the non-linear noise in long-haul DWDM transmission. Several proprietary and open-source tools use some form of GN models in order to calculate the SNR for the physical path validity of optical services. GN models have both analytical as well as closed-form solutions, which are derived by simplifying some assumptions. However, all the models follow the two basic assumptions, which form the crux of all GN models. These assumptions are briefly described as follows:

1. Signal Gaussianity - In this assumption, the transmitted signal can be modeled as a circular stationary gaussian noise.
2. NLIN Independence - In this assumption, the NLIN at the output is gaussian and completely independent from the input power or ASE noise.
3. Flat PSD - The PSD of the NLIN is assumed flat for a single channel bandwidth (i.e. no roll-off).

Since GN model equations are both highly complex, as well as out of the scope of this dissertation, we point interested readers to the equations from relevant literature which have been implemented in the path calculation engines used in this dissertation.

A.1 GN Reference Formula

The GN reference formula was featured in Poggiolini *et. al.* (Eq. 1) [60]. It provides a general model, which is capable of describing different types of DWDM systems, including both

lumped (e.g. EDFA) gain amplification and distributed field (e.g. Raman) gain amplification. Since Eq. 1 in Ref. [60] is general, it is assumed that all span lengths in the optical link are equal and that the amplifier at the end of the span perfectly compensates for the loss of the span. Extending the four-wave mixing definition, the GN reference formula defines the NLIN PSD at a given frequency f as the integrated sum of all infinitesimally small four-wave mixing products among any three frequencies in the spectrum, namely, $f1$, $f2$, and $f3$. Here $f3$ is a perturbative result of the interaction between $f1$, $f2$, and f. It is defined as $f3 = f1 + f2 - f$.

A.2 Enhanced Gaussian Noise Model

Following the deployment and experimental validation of the GN reference formula and its approximations, it was observed that the model was overestimating the NLIN in the first spans of the link. Therefore, the mitigation of these estimation errors, along with a new set of equations to define self-channel, cross-channel and multi-channel interference was defined by Carena *et. al.* [59] and termed as Enhanced Gaussian Noise Model (EGN) model. Equations 5 and 18 of this work refer to the SCI and XCI, which have been implemented in this dissertation. Compared to the approximations of the GN reference formula, the EGN model reduces the overestimation by almost 5 dB when the number of spans in the link is smaller. This is also observed consistently as the number of channels in the system is varied. However, the equations are rather complex to deploy on a fully loaded DWDM link with many channels, leading to extended NLIN calculation times.

A.3 Accurate Closed Form Enhanced GN Model

For C-band DWDM transmission systems, the EGN model is not a practical tool, since higher calculation time leads to long waiting times for network planning results. For online network planning applications, like physical path calculation within a network, approximations of the full-form EGN model enable us to achieve faster simulation times, while maintaining the required accuracy. The Accurate Closed Form Enhanced Gaussian Noise Model (ACF-EGN) model defined by Ranjbar-Zafreh *et. al.* is one such approximation that has been implemented in this dissertation [61]. The approximation depends on several parameters, which were first optimized for only fully-loaded C-band systems. However, later, these parameters were optimized using machine-learning techniques (out of the scope of this dissertation). This model was able to replicate the results of the full-form EGN model with an accuracy of 0.2 dB, thereby allowing it to be implemented in physical path-aware network planning tools.

A.4 Inter-channel Stimulated Raman Scattering GN model

The previously implemented models focussed on NLIN estimation purely on C-band systems. However, as DWDM transmission experiments in L-band and S-band were successfully demonstrated, there was a need for a GN model that can accurately estimate NLIN in more than C-band transmission systems. For this, the Inter-Sysmbol Stimulated Raman Scattering Gaussian Noise (ISRS-GN) model presented by Semrau *et. al.* [33] has been implemented in this dissertation to calculate the NLIN for network planning studies in Chapter 4. The ISRS-GN model takes into account frequency-dependent attenuation along the fiber, as well as the effects of the Raman gain over the entire wideband system. The ISRS-GN model changes the NLIN of a fully-loaded C+L-band system at optimum launch power by 2 dB, thereby proving the importance of considering stimulated Raman scattering effects in C+L-band and C+L+S-band systems.

Appendix B

Selected Machine-Learning Models

Here we briefly discuss the theoretical aspects of the machine learning models used in this dissertation. In Chapter 4, reinforcement learning was used, whereas in Chapter 5, several ML and statistical algorithms were implemented for failure detection. These are now discussed briefly here. For a complete evaluation of reinforcement learning strategies of multi-band network planning and failure detection strategies, interested readers can refer to the master theses of Arantxa Paz [144], and Isabella Dick [145] respectively.

B.1 Reinforcement Learning

Reinforcement Learning (RL) is an area of machine learning where agents are created to be trained by experience. The agents make decisions and interact with an environment, this way they learn which are the best decisions and which not. The objective is to maximize a reward, parameter that expresses a long-term benefit. RL in this way, tries to mimic the experience of human learning through trial and error [146]. In Figure B.1, the components of RL are presented. There is an **environment**, which is the interface to interact with. The **agent** is the one that makes decisions. When an agent makes a decision, the environment transitions into a new **state**. A **reward** is an indicator that tells if the action taken is correct and it is generally numeric [147]. The utility of RL is broad and promising for many industries and is focused on optical network planning in this dissertation.

B.1.1 Markov Decision Process

To explain the Markov Decision Process (MDP), a review of Markov chains is necessary. Markov chains model stochastic processes. A state is information that describes the current situation of an environment. The processes have the Markov property when the transition from one state to another depends only on the current state and not the past.

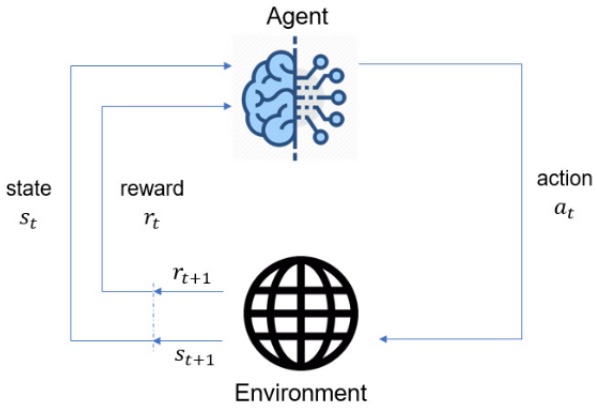


Figure B.1: Reinforcement Learning components [146].

Markov Decision Process (MDP) presents the concept of reinforcement learning in a mathematical form. An agent takes an action and influences the environment. The goal is to achieve the maximum reward by choosing a set of actions.

B.1.2 Q-learning

Q-learning is a very popular algorithm in reinforcement learning. Equation B.1 shows its online update rule, where online means that the policy update is immediate and the next action is better than the previous. Different from SARSA, in Q learning, there are two policies. One is the epsilon-greedy policy, and the other is a greedy policy. The action is selected using an epsilon-greedy policy, while the Q value of the next state-action pair uses a greedy policy. Since a greedy policy always selects the action with the maximum value, the term *max* can be introduced in Equation B.1 [146, 148, 149].

Broadly, Q-learning falls under the category of an off-policy learning algorithm. An off-policy algorithm uses two policies, one for selecting an action in the environment and the other to find the optimal policy [149]. Both algorithms initialize a Q function with random values or zeros. They extract a policy from $Q(s,a)$ and select action a for state s. Then, for each step in each episode, they perform an action a and move to the next state s', observing reward r. In state s', new action a' is selected, and $Q(s,a)$ is updated with the corresponding equation (Eq. B.1). The s and a values are updated with the new state-action $(s'-a')$ pair. If not in a terminal state, a new step is iterated. Hence, the difference between the two algorithms is only in the update of the Q value [149].

$$Q(s,a) \longleftarrow Q(s,a) + \alpha(r + \gamma \max_{a} Q(s',a') - Q(s,a)) \tag{B.1}$$

B.1.3 Deep Q-learning Network

A Deep Q-Network (DQN) arises from using a deep neural network to achieve Q-learning tasks. Although this concept was available for many decades, it has been possible only recently to implement it, due to its high computational requirements. In a DQN, the states are mapped as an input to the deep neural network, and the expected value of actions is mapped as the output. Hence, the state spaces do not need to be discrete and can handle complex features [147]. DQN has the following three parts which need to be implemented:

- Experience Replay: It is a buffer with information from observations, actions, and rewards. When training the model, a subset of the buffer is taken which helps to avoid correlations between contiguous observations. The size of the experience replay buffer can greatly affect the performance, as it needs to be big enough to not "forget" important data. However, larger replay buffers can also lead to slower training and prediction.
- Q-Network clones: There are two neural networks in a DQN namely, an online network that produces actions, and a target network that is assigned for learning. After certain iterations, the weights from the target network will be copied to the prediction network. This period must be small to improve the model fast but not too small to keep the data stationary.
- Choice of Neural Network: The neural networks map the observations to actions. Hence, depending on the observation, an appropriate architecture (e.g., ANN, Convolutional Neural Network (CNN), etc.) and activation function of the neural network must be chosen.

B.2 Models for Failure Detection

Failure detection is implemented using four different methods, namely OCSVM, ANN, ORP prediction with dynamic threshold, and ARIMAX. Of these four, ORP prediction with dynamic threshold and ARIMAX do not use ML-based models for their predictions. Further, the comparisons between the ML and non-ML approaches are done in order to find explainable solutions to the problem of failure detection for OSaaS users.

B.2.1 One-Class Support Vector Machine (OCSVM)

Support Vector Machine (SVM) is a classification algorithm, which is a subset of supervised learning methods. It was originally invented to perform binary classification. The method

is based on a set of labeled training objects, where each object is represented by a vector $\mathbf{x}_i$ labeled by y_i,

$$(\mathbf{x}_i, y_i) \forall i \in \{1, m\}, y \in \{-1, 1\} \tag{B.2}$$

where m is the number of training objects. The SVM aims to find a hyperplane such that it divides the training objects into two classes. The distance between the hyperplane and the vectors that are closest to the hyperplane, also called support vectors, is maximized. The hyperplane is described as:

$$(\mathbf{w}, \mathbf{x}) + b = 0 \tag{B.3}$$

where the normal vector $\mathbf{w}$ represents a straight line through the coordinate origin. Perpendicular to this vector are hyperplanes, where each intersects the vector at a certain bias distance $b/||\mathbf{w}||^2$ to the origin. The combination of the normal vector and the bias b uniquely determine a hyperplane, and for all points x belonging to it, Eq. B.2 holds. For points, that are not on the hyperplane, the value is either positive or negative. The sign defines on which side of the hyperplane the point lies. If a class membership in the training examples is expressed by $y_i = \pm 1$, the class of a vector can be obtained by the formulaic condition,

$$y_i = \text{sign}((\mathbf{w}, \mathbf{x_i}) + b) \tag{B.4}$$

Hyperparameters, that can be adjusted include the kernel function (linear, polynomial with degree, sigmoid, or radial basis) and the parameter ν, which defines the upper bound on the fraction of training errors as well as a lower bound of the fraction of support vectors. It can also be seen as the ratio of outlier samples in the dataset.

B.2.2 Autoregressive Integrated Moving Average with Exogenous Variables (ARIMAX)

A very common method in time series forecasting is the Autoregressive Integrated Moving Average (ARIMA) model, which aims to describe the autocorrelations in the data. It is fitted to time series data to predict future points in the series. The prediction is parameterized by three distinct integers:

- p: Number of autoregressive terms, incorporating the effect of past values into the model (autoregressive term)
- d: Number of nonseasonal differences needed for stationarity (integrated term)
- q: Number of lagged forecast errors in the prediction equation (moving average term)

Mathematically, the model can be described as followed:

$$y(t) = \alpha + \sum_{i=1}^{p} \phi_i y(t-1) + \sum_{j=1}^{q} \sigma_j s(t-j) + s(t) \tag{B.5}$$

where α is a constant, the ϕ terms denote linear combination lags of y (up to p lags) and the σ terms represent a linear combination of lagged forecast errors s (up to q lags). It is possible to extend the ARIMA model by using exogenous variables in order to fit multi-variate data. In literature, the algorithm is often named ARIMAX and is defined as:

$$y(t) = \alpha + \sum_{i=1}^{p} \phi_i y(t-1) + \sum_{j=1}^{q} \sigma_j s(t-j) + s(t) + \sum_{m=1}^{b} \eta_m x(t-m) \tag{B.6}$$

where $\eta_1, ..., \eta_b$ are the parameters of the exogenous input x(t) [112]. Since the underlying OPM data is a multivariate time series, ARIMAX is used to predict the next value of the received power. Thus, p, q, and d are found for the best fit of the algorithm.

Appendix C

Algorithms for Network Planning and Upgrade

In this section, we define the workflow of the combinatorial optimization based multi-period multi-band planning algorithms used in Chapter 4.

C.1 RCSA Algorithms

Algorithm 1 provides the pseudo-code to first route a candidate lightpath for a given demand using Yen's k-shortest path algorithm, and then order them from longest to shortest path lengths. Then, for each demand, given the path length, a list of possible configurations can be filtered using the HeCSON algorithm (ref. Sec. 3.1). Algorithm 2 provides the pseudo-code for multi-period RCSA algorithm. Given a network and a set of routed demands as well as configurations for each demand, this algorithm first tries to upgrade existing lightpaths in the network, and then places new lightpaths to meet the traffic in each planning year.

C.2 Migration Strategy Algorithms

The band and fiber upgrade strategies based on reinforcement learning, discussed in Section 4.3 are shown as a pseudo-code here. Algorithm 5 refers to the pseudo-code for conducting link upgrade action in the network using reinforcement learning. Similarly, Algorithm 6 refers to the reward updation process after each time the action process is executed. Algorithms 3 and 4 define the two functions $reserveFromExisting$ and $reserveFromAny$ used in Algorithm 5 during the action process. The definitions of the variables and parameters are given in Table 4.5.

Algorithm 1: Routing and Configuration pre-selection algorithm.

function RouteAndPreselect ($\mathcal{G}(\mathcal{V},\mathcal{E},\mathcal{D})$, C);
Input : Network topology $\mathcal{G}(\mathcal{V},\mathcal{E},\mathcal{D})$ and BVT/CFP2-DCO configuration list C
Output Ordered list of routed demands $\mathcal{D}$ and possible configurations for all demands $C_{\mathcal{D}}$
:
for $d \in \mathcal{D}$ **do**
  $SP_d \leftarrow$ Yen's k = 3 shortest paths of d;
  $Hops_d \leftarrow$ num hops of $SP_d[0]$;
end
Order $d \in \mathcal{D}$ longest to shortest SP_d then highest to lowest $Hops_d$;
for $d \in \mathcal{D}$ **do**
  for $k \leftarrow 0$ *to* 2 **do**
    for $c \in C$ **do**
      Place set of lightpaths $L_{d,k,c}$ with config c on all freq slots of links $e \in SP_d$;
      Calculate $GSNR_{d,k,c,l} \forall\, l \in L_{d,k,c}$ [33] ;
      $GSNR_{d,k,c} \leftarrow \min(GSNR_{d,k,c,l})$;
      if $GSNR_{d,k,c} \geq minSNR_c$ **then**
        $c \rightarrow C_{d,k}$;
      end
      Remove $L_{d,k,c}$ from all freq slots ;
    end
    $C_{d,k} \rightarrow C_d$;
  end
  $C_d \rightarrow C_{\mathcal{D}}$;
end
return $\mathcal{D}$, $C_{\mathcal{D}}$

Algorithm 2: Multi-period RCSA Algorithm.

function MultiperiodRCSA ($\mathcal{G}(\mathcal{V},\mathcal{E},\mathcal{D})$, C, T);
Input : Network topology $\mathcal{G}(\mathcal{V},\mathcal{E},\mathcal{D})$, transceiver configuration list C, total planning period T
Output List of placed lightpaths $\mathcal{L}$, yearly underprovisioning $\mathcal{U}$
:
$\mathcal{D}$, $C_{\mathcal{D}} \leftarrow$ RouteAndPreselect($\mathcal{G}(\mathcal{V},\mathcal{E},\mathcal{D})$, C) [Alg. 1]
for $t \in \mathcal{T}$ **do**
  for $d \in \mathcal{D}$ **do**
    $(i,j) \leftarrow d$;
    $\delta(i,j,t) \leftarrow \tau_t(d)$;
    $\widetilde{L_d} \leftarrow$ UpgradeLP($L_d \in \mathcal{L}$) [4];
    Recalculate GSNR for all placed lightpaths sharing the same links as $\widetilde{L_d}$;
    Calculate $\rho(i,j,t)$;
    if $\rho(i,j,t) \geq 0$ **then**
      $L_{d,t} \leftarrow$Eqs. 4.3-4.5;
    else
      continue;
    end
    $L"_{d,t} \leftarrow$ First-Fit($L_{d,t}$);
    if $\rho(i,j,t) \geq \sum L"_{d,t}$ **then**
      $L"_{d,t} \leftarrow$ Reroute($d,\mathcal{L},t$) [85]
    end
    Recalculate GSNR for all placed lightpaths sharing the same links as $L"_{d,t}$;
    $\rho(i,j,t) \leftarrow \rho(i,j,t) - \sum_{\forall l \in L"_{d,t}} DR_l$;
    $L"_{d,t} \rightarrow \mathcal{L}$;
    if $\rho(i,j,t) \geq 0$ **then**
      Increase underprovisioning $\mathcal{U}$ using Eq. 4.2
    end
  end
end
return $\mathcal{L}$, $\mathcal{U}$

Algorithm 3: reserveFromExisting

Data: $d, path, band, fiber, Slots^{d}_{p,b,f}, Stock^{d}_{p,b,f}, A^{d}_{p,b,f}$
Result: $path, band, fiber$

```
1  for p ∈ P do
2      for f ∈ F do
3          for b ∈ B do
4              if Slots^d_{p,b,f} <= Stock^d_{p,b,f} and A^d_{path,b,f} == 1 then
5                  if (B × f + b) < (B × fiber + band) then
6                      path = p; band = b; fiber = f;
7                      break
8                  end
9              end
10         end
11         if band!=1000 then
12             break
13         end
14     end
15 end
```

Algorithm 4: Function reserveFromAny

Data: $d, path, band, fiber, Slots^{d}_{p,b,f}, Stock^{d}_{p,b,f}, A^{d}_{p,b,f}$
Result: $path, band, fiber$

```
1  for p ∈ P do
2      for f ∈ F do
3          for b ∈ B do
4              if Slots^d_{p,b,f} <= Stock^d_{p,b,f} and A^d_{path,b,f} == 1 then
5                  if (B × f + b − Σ_{B,F} A^d_{p,b',f'}) < (B × fiber + band − Σ_{B,F} A^d_{path,b',f'}) then
6                      path=p;
7                      band=b;
8                      fiber=f;
9                      break
10                 else if (B × f + b − Σ_{B,F} A^d_{p,b',f'}) == (B × fiber + band − Σ_{B,F} A^d_{path,b',f'}) and
                       ((B × f + b) < (B × fiber + band)) then
11                     path=p;
12                     band=b;
13                     fiber=f;
14                     break
15             end
16         end
17         if band!=1000 then
18             break
19         end
20     end
21 end
```

Algorithm 5: RL solution ActionProcess (for a demand in a given period)

Data: $action, d, Slots^{d}_{p,b,f}, Stock^{d}_{p,b,f}, A^{d}_{p,b,f}, GeneralStock^{l}_{b,f}, GeneralActivation^{l}_{b,f}, X^{d}_{p,l}$
Result: $Reward, GeneralStock^{l}_{b,f}, GeneralActivation^{l}_{b,f}$

$path, band, fiber \leftarrow 1000$
$path, band, fiber \leftarrow$reserveFromExisting$(d, path, band, fiber, Slots^{d}_{p,b,f}, Stock^{d}_{p,b,f}, A^{d}_{p,b,f})$
if *band is 1000* **then**
  $path, band, fiber \leftarrow$ reserveFromAny$(d, path, band, fiber, Slots^{d}_{p,b,f}, A^{d}_{p,b,f})$
  if *band is 1000* **then**
    Print 'Not enough assigned fibers'
  end
end
$Status \leftarrow \sum_{b,f} A^{d}_{path,b,f}$ /* Expected threshold for no upgrade */
$Chosen \leftarrow band + fiber * band + 1$ /* Chosen number */
if $action >= 1$ *and* $action <= 3$ **then**
  $GeneralActivation^{l}_{b,f} \leftarrow$ upgradeOneBand$(path, GeneralActivation^{l}_{b,f}, X^{d}_{p,l})$
end
$Reward, GeneralStock^{l}_{b,f} \leftarrow getReward(action, d, path, band, fiber, Slots^{d}_{p,b,f}, GeneralStock^{l}_{b,f})$

Algorithm 6: RL solutionn getReward (for a demand in a given period)

Data: $action, d, band, fiber, path, Slots^{d}_{p,b,f}, GeneralStock^{l}_{b,f}$
Result: $Reward, GeneralStock^{l}_{b,f}$

if $Chosen - Status <= 0$ **then**
  if $Action = 0$ **then**
    $Reward \leftarrow 1$
  else
    $Reward \leftarrow -1$
  end
  $GeneralStock^{l}_{b,f} \leftarrow$ provisionDemand$(Slots^{d}_{p,b,f}, band, fiber, path, d)$
else if $Chosen - Status = 1$ **then**
  if $Action == (path + 1)$ **then**
    $Reward \leftarrow 1$
    $GeneralStock^{l}_{b,f} \leftarrow$ provisionDemand$(Slots^{d}_{p,b,f}, band, fiber, path, d)$
  else if *Action!=0* **then**
    $Reward \leftarrow -1$
    $GeneralStock^{l}_{b,f} \leftarrow$ provisionDemand$(Slots^{d}_{p,b,f}, band, fiber, path, d)$
  else
    $Reward \leftarrow -1$
  end
else
  Print 'An upgrade higher than two bands is needed'
  $Reward \leftarrow 0$
end

Bibliography

Publications by the author

Journal publications

[1] S. K. Patri, A. Autenrieth, J.-P. Elbers, and C. Mas-Machuca. "Multi-band transparent optical network planning strategies for 6G-ready European networks." In: *Optical Fiber Technology* 74 (2022), p. 103118.

[2] S. K. Patri, A. Villavicencio Paz, M. Wenning, J. Müller, A. Autenrieth, et al. "Reinforcement Learning-enabled Optical Transport Network Migration Strategies [SUBMITTED]." In: *Journal of Optical Communications and Networking* xx.x (2023), pp. x–x.

Conference publications

[3] S. K. Patri, A. Autenrieth, D. Rafique, J.-P. Elbers, and C. Mas-Machuca. "HeCSON: heuristic for configuration selection in optical network planning." In: *Optical Fiber Communication Conference*. Optica Publishing Group. 2020, Th2A–32.

[4] S. K. Patri, A. Autenrieth, J.-P. Elbers, and C. Mas-Machuca. "Planning optical networks for unexpected traffic growth." In: *2020 European Conference on Optical Communications (ECOC)*. IEEE. 2020, pp. 1–4.

[5] D. Khomchenko, S. K. Patri, A. Autenrieth, C. Mas-Machuca, and A. Richter. "Transmission-Aware Bandwidth Variable Transceiver Allocation in DWDM Optical Networks." In: *2021 International Conference on Optical Network Design and Modeling (ONDM)*. IEEE. 2021, pp. 1–6.

[6] C. Mas-Machuca, S. K. Patri, and S. Amjad. "Long-term Capacity Planning in Flexible Optical Transport Networks." In: *2022 Optical Fiber Communications Conference and Exhibition (OFC)*. IEEE. 2022, pp. 1–3.

[7] S. K. Patri, I. Dick, K. Kaeval, J.-J. Pedreno-Manresa, A. Autenrieth, et al. “Machine Learning enabled Fault-Detection Algorithms for Optical Spectrum as a Service Users [ACCEPTED].” In: *27th International Conference on Optical Network Design and Modelling (ONDM)*. IEEE. 2023. to appear.

[8] S. K. Patri, S. Sultana, M. Dürre, S. Amjad, A. Schumann, et al. “Open-Source Service Management for a Fully Disaggregated Optical Network Simulation.” In: *2022 18th International Conference on Network and Service Management (CNSM)*. IEEE. 2022, pp. 364–366.

[9] S. K. Patri, M. Wenning, S. H. Gonde, A. Autenrieth, J.-P. Elbers, et al. “Trusted Node Deployment Strategies for Long-Haul Quantum Key Distribution Networks [ACCEPTED].” In: *27th International Conference on Optical Network Design and Modelling (ONDM)*. IEEE. 2023. to appear.

[10] J. Y. Cho, S. K. Patri, and A. Sergeev. “Demonstration of Flexible and Scalable Quantum-Resistant Encryption with Threshold Key Management in Optical Networks.” In: *2020 European Conference on Optical Communications (ECOC)*. IEEE. 2020, pp. 1–4.

[11] J. Y. Cho, J.-J. Pedreno-Manresa, S. K. Patri, A. Sergeev, J.-P. Elbers, et al. “Demonstration of Software-defined Key Management for Quantum Key Distribution Network.” In: *Optical Fiber Communication Conference*. Optical Society of America. 2021, M2B–4.

Others

[12] *Physical Network Information*. Accessed: Aug 2020. URL: www.github.com/SaiPatri/PhyNWInfo.

General publications

[13] M. Ruiz, L. Velasco, A. Lord, D. Fonseca, M. Pioro, et al. "Planning fixed to flexgrid gradual migration: drivers and open issues." In: *IEEE Communications Magazine* 52.1 (2014), pp. 70–76. DOI: 10.1109/MCOM.2014.6710066.

[14] E. Commission. *Upcoming report on the rollout of 5G in the EU*. Accessed: April 2022. URL: https://www.eca.europa.eu/en/Pages/NewsItem.aspx?nid=16334.

[15] D. Goovaerts. *Altice USA loses 13K broadband subs as it waits for fiber growth*. Accessed: April 2022. URL: https://www.fiercetelecom.com/broadband/altice-usa-loses-13k-broadband-subs-it-waits-fiber-growth.

[16] *G.694.1 : Spectral grids for WDM applications: DWDM frequency grid*. https://www.itu.int/rec/T-REC-G.694.1/, Accessed: Jan 2023.

[17] K. C. Kao and G. A. Hockham. "Dielectric-fibre surface waveguides for optical frequencies." In: *Proceedings of the Institution of Electrical Engineers*. Vol. 113. 7. IET. 1966, pp. 1151–1158.

[18] *ADVA Teraflex*. Accessed: Aug 2020. URL: www.adva.com/en/products/open-optical-transport/fsp-3000-open-terminals/teraflex.

[19] Acacia. *CFP2-DCO Product Family*. https://acacia-inc.com/product/cfp2/, Accessed: April 2022.

[20] A. Ferrari, E. Virgillito, and V. Curri. "Band-Division vs. Space-Division Multiplexing: A Network Performance Statistical Assessment." In: *Journal of Lightwave Technology* 38.5 (2020), pp. 1041–1049. DOI: 10.1109/JLT.2020.2970484.

[21] B. Shariati, J. M. Rivas-Moscoso, D. Klonidis, I. Tomkos, S. Ben-Ezra, et al. "Options for Cost-effective Capacity Upgrades in Backbone Optical Networks." In: *2016 21st European Conference on Networks and Optical Communications (NOC)*. 2016, pp. 35–40. DOI: 10.1109/NOC.2016.7506982.

[22] V. López, L. Velasco, et al. "Elastic optical networks." In: *Architectures, Technologies, and Control, Switzerland: Springer Int. Publishing* (2016).

[23] *ONF Open Disaggregated Transport Network (ODTN)*. Accessed: Feb 2023. URL: https://opennetworking.org/odtn/.

[24] *OpenDaylight TransportPCE User Guide*. Accessed: Feb 2023. URL: https://docs.opendaylight.org/projects/transportpce/en/latest/user-guide.html.

[25] A. Ferrari, M. Filer, E. Le Rouzic, J. Kundrát, B. Correia, et al. "GNPy: an open source planning tool for open optical networks." In: *2020 International Conference on Optical Network Design and Modeling (ONDM)*. 2020, pp. 1–6. DOI: 10.23919/ONDM48393.2020.9133027.

[26] A. Giorgetti, A. Sgambelluri, R. Casellas, R. Morro, A. Campanella, et al. "Control of open and disaggregated transport networks using the Open Network Operating System (ONOS) [Invited]." In: *Journal of Optical Communications and Networking* 12.2 (2020), A171–A181. DOI: 10.1364/JOCN.12.00A171.

[27] *Quantum Key Distribution (QKD); QKD Module Security Specification*. https://www.etsi.org/deliver/etsi_gs/qkd/001_099/008/01.01.01_60/gs_qkd008v010101p.pdf, Accessed: Jan 2023.

[28] *Quantum Key Distribution Networks - Key Management*. https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-Y.3803-202012-I!!PDF-E&type=items, Accessed: Jan 2023.

[29] *ID Quantique | Quantum Key Distribution | Cerberis XG QKD System*. https://www.idquantique.com/quantum-safe-security/products/cerberis-xg-qkd-system/. (Accessed on 02/27/2023).

[30] Z. I. Berlin. *SNDLib*. Accessed: November 2020. URL: https://sndlib.zib.de.

[31] G. He, S. Searcy, D. Gariépy, and S. Tibuleac. "GOSNR Characterization by Optical Spectrum Analysis." In: *2020 Optical Fiber Communications Conference and Exhibition (OFC)*. 2020, pp. 1–3.

[32] P. C. Becker, N. A. Olsson, and J. R. Simpson. *Erbium-Doped Fiber Amplifiers: Fundamentals and Technology*. 1st ed. Vol. 1. Academic Press, 1999.

[33] D. Semrau, L. Galdino, E. Sillekens, D. Lavery, R. I. Killey, et al. "*Modulation format dependent, closed-form formula for estimating nonlinear interference in S+C+L band systems*." In: *45th European Conference on Optical Communication (ECOC 2019)*. 2019, pp. 1–4. DOI: 10.1049/cp.2019.0892.

[34] J. L. Zyskind, J. A. Nagel, and H. D. Kidorf. "Erbium-doped fiber amplifiers for optical communications." In: *Optical Fiber Telecommunications IIIB* (1997), pp. 13–68.

[35] B. T. Teipen, M. H. Eiselt, K. Grobe, and J.-P. Elbers. "Adaptive data rates for flexible transceivers in optical networks." In: *Journal of Networks* 7.5 (2012), p. 776.

[36] *Internet Topology Zoo*. http://www.topology-zoo.org/dataset.html, Accessed: April 2022.

[37] B. Correia, R. Sadeghi, E. Virgillito, A. Napoli, N. Costa, et al. “Networking Performance of Power Optimized C+L+S Multiband Transmission.” In: *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*. 2020, pp. 1–6. DOI: 10.1109/GLOBECOM42002.2020.9322068.

[38] A. Dwivedi et al. “Traffic model for USA long-distance optical network.” In: *Proc. OFC*. 2000.

[39] *Germany DE-CIX*. Accessed: Aug 2020. URL: www.de-cix.net/en/locations/germany.

[40] T. Ahmed, S. Rahman, M. Tornatore, X. Yu, K. Kim, et al. “Dynamic Routing and Spectrum Assignment in Co-Existing Fixed/Flex-Grid Optical Networks.” In: *2018 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*. 2018, pp. 1–3. DOI: 10.1109/ANTS.2018.8710140.

[41] E. E. Moghaddam, H. Beyranvand, and J. A. Salehi. “Routing, Spectrum and Modulation Level Assignment, and Scheduling in Survivable Elastic Optical Networks Supporting Multi-Class Traffic.” In: *Journal of Lightwave Technology* 36.23 (2018), pp. 5451–5461. DOI: 10.1109/JLT.2018.2874820.

[42] G. Choudhury, D. Lynch, G. Thakur, and S. Tse. “Two use cases of machine learning for SDN-enabled IP/optical networks: traffic matrix prediction and optical path performance prediction [Invited].” In: *Journal of Optical Communications and Networking* 10.10 (2018), pp. D52–D62. DOI: 10.1364/JOCN.10.000D52.

[43] Cisco. *Cisco Network Convergence System 1004 Data Sheet*. Accessed: November 2020. URL: https://www.cisco.com/c/en/us/products/collateral/optical-networking/network-convergence-system-1000-series/datasheet-c78-740368.html.

[44] O. Turkcu et al. “Robust Optical Networks with Emerging Coherent Technologies and Traffic Trends.” In: *Advanced Photonics*. 2018.

[45] *CISCO VNI Global Forecast 2021*. https://www.cisco.com/c/dam/m/en_us/solutions/service-provider/vni-forecast-highlights/pdf/Global_2021_Forecast_Highlights.pdf. (Accessed on 02/27/2023).

[46] A. Soule, A. Nucci, R. Cruz, E. Leonardi, and N. Taft. “How to identify and estimate the largest traffic matrix elements in a dynamic environment.” In: *Proceedings of the joint international conference on Measurement and modeling of computer systems*. 2004, pp. 73–84.

[47] A. Dwivedi and R. Wagner. "Traffic model for USA long-distance optical network." In: *Optical Fiber Communication Conference. Technical Digest Postconference Edition. Trends in Optics and Photonics Vol.37 (IEEE Cat. No. 00CH37079)*. Vol. 1. 2000, 156–158 vol.1. DOI: 10.1109/OFC.2000.868400.

[48] V. Lopez, B. Zhu, D. Moniz, N. Costa, J. Pedro, et al. "Optimized Design and Challenges for C&L Band Optical Line Systems." In: *Journal of Lightwave Technology* 38.5 (2020), pp. 1080–1091. DOI: 10.1109/JLT.2020.2968225.

[49] M. Bjorklund. *The YANG 1.1 data modeling language*. Tech. rep. 2016.

[50] R. Enns. *NETCONF Configuration Protocol*. RFC 4741. Accessed: February 2023. RFC Editor, Dec. 2006. URL: http://www.rfc-editor.org/rfc/rfc4741.txt.

[51] OpenROADM. *OpenROADM Multi Source Agreement*. http://openroadm.org/. Accessed: July 2022.

[52] *Products | Quantum Key Distribution | TOSHIBA DIGITAL SOLUTIONS CORPORATION*. https://www.global.toshiba/ww/products-solutions/security-ict/qkd/products.html. (Accessed on 02/27/2023).

[53] C. H. Bennett and G. Brassard. "An update on quantum cryptography." In: *Workshop on the theory and application of cryptographic techniques*. Springer. 1985, pp. 475–480.

[54] X. Tang, L. Ma, A. Mink, A. Nakassis, H. Xu, et al. "Quantum key distribution system operating at sifted-key rate over 4 Mbit/s." In: *Quantum Information and Computation IV*. Vol. 6244. SPIE. 2006, pp. 182–189.

[55] E. Diamanti. "Security and implementation of differential phase shift quantum key distribution systems." PhD thesis. Stanford University, Jan. 2006.

[56] X. Ma. "Quantum cryptography: theory and practice." In: *arXiv preprint:0808.1385* (2008).

[57] *TeraFlex™*. https://www.adva.com/en/products/open-optical-transport/fsp-3000-open-terminals/teraflex. (Accessed on 02/27/2023).

[58] J.-L. Auge, G. Grammel, E. le Rouzic, V. Curri, G. Galimberti, et al. "Open Optical Network Planning Demonstration." In: *2019 Optical Fiber Communications Conference and Exhibition (OFC)*. 2019, pp. 1–3.

[59] A. Carena, G. Bosco, V. Curri, Y. Jiang, P. Poggiolini, et al. "EGN model of non-linear fiber propagation." In: *Optics express* 22.13 (2014), pp. 16335–16362.

[60] P. Poggiolini, G. Bosco, A. Carena, V. Curri, Y. Jiang, et al. “A detailed analytical derivation of the GN model of non-linear interference in coherent optical transmission systems.” In: *arXiv preprint arXiv:1209.0394* (2012).

[61] M. R. Zefreh, A. Carena, F. Forghieri, S. Piciaccia, and P. Poggiolini. “Accurate closed-form gn/egn-model formula leveraging a large qam-system test-set.” In: *IEEE Photonics Technology Letters* 31.16 (2019), pp. 1381–1384.

[62] P. Poggiolini et al. “Accurate Non-Linearity Fully-Closed-Form Formula based on the GN/EGN Model and Large-Data-Set Fitting.” In: *Proc. OFC*. 2019.

[63] *Network Optimization in the 600G Era*. https://acacia-inc.com/wp-content/uploads/2018/12/Network-Optimization-in-the-600G-Era-WP1218.pdf. (Accessed on 02/27/2023).

[64] *Keeping the Internet up and running in times of crisis*. Accessed: Aug 2020. URL: www.oecd.org/coronavirus/policy-responses/keeping-the-internet-up-and-running-in-times-of-crisis-4017c4c9.

[65] R. Wang et al. “Load-aware Nonlinearity Estimation for Elastic Optical Network Resource Optimization and Management.” In: *IEEE/OSA JOCN* 11.5 (2019).

[66] E. Virgillito et al. “Statistical Assessment of Open Optical Networks.” In: *MDPI Photonics* 6.2 (2019).

[67] L. Velasco, A. Castro, M. Ruiz, and G. Junyent. “Solving Routing and Spectrum Allocation Related Optimization Problems: From Off-Line to In-Operation Flexgrid Network Planning.” In: *Journal of Lightwave Technology* 32.16 (2014), pp. 2780–2795. DOI: 10.1109/JLT.2014.2315041.

[68] P. Soumplis, K. Christodoulopoulos, M. Quagliotti, A. Pagano, and E. Varvarigos. “Network planning with actual margins.” In: *Journal of Lightwave Technology* 35.23 (2017), pp. 5105–5120.

[69] X. Chen, J. Guo, Z. Zhu, R. Proietti, A. Castro, et al. “Deep-RMSA: A Deep-Reinforcement Learning Routing, Modulation and Spectrum Assignment Agent for Elastic Optical Networks.” In: *2018 Optical Fiber Communications Conference and Exposition (OFC)*. 2018, pp. 1–3.

[70] C. Natalino and P. Monti. “The Optical RL-Gym: An open-source toolkit for applying reinforcement learning in optical networks.” In: *2020 22nd International Conference on Transparent Optical Networks (ICTON)*. 2020, pp. 1–5. DOI: 10.1109/ICTON51198.2020.9203239.

[71] X. Zhou, W. Lu, L. Gong, and Z. Zhu. "Dynamic RMSA in elastic optical networks with an adaptive genetic algorithm." In: *2012 IEEE Global Communications Conference (GLOBECOM)*. 2012, pp. 2912–2917. DOI: 10.1109/GLOCOM.2012.6503559.

[72] J. M. Simmons. *Optical network design and planning*. Springer, 2014.

[73] K. Walkowiak, M. Klinkowski, A. Włodarczyk, and A. Kasprzak. "Predeployment of transponders for dynamic lightpath provisioning in translucent spectrally–spatially flexible optical networks." In: *Applied Sciences* 10.8 (2020), p. 2802.

[74] M. A. Cavalcante, H. A. Pereira, D. A. Chaves, and R. C. Almeida. "Evolutionary multiobjective strategy for regenerator placement in elastic optical networks." In: *IEEE Transactions on Communications* 66.8 (2018), pp. 3583–3596.

[75] T. Zami, A. Morea, and J. Pesic. "Benefit of progressive deployment of regenerators along with traffic growth in WDM elastic networks." In: *2018 Optical Fiber Communications Conference and Exposition (OFC)*. IEEE. 2018, pp. 1–3.

[76] P. Soumplis, K. Christodoulopoulos, M. Quagliotti, A. Pagano, and E. Varvarigos. "Multi-period planning with actual physical and traffic conditions." In: *Journal of Optical Communications and Networking* 10.1 (2018), A144–A153.

[77] F. Lezama, A. F. Martínez-Herrera, G. Castañón, C. Del-Valle-Soto, A. M. Sarmiento, et al. "Solving routing and spectrum allocation problems in flexgrid optical networks using pre-computing strategies." In: *Photonic Network Communications* 41.1 (2021), pp. 17–35.

[78] S. S. Skiena. *The algorithm design manual*. Vol. 2. Springer, 1998.

[79] D. Khomchenko, D. Yevseyenko, J. Farina, and A. Richter. "Automated design of add/drop equipment and effective wavelength assignment in complex DWDM networks." In: *Metro and Data Center Optical Networks and Short-Reach Links*. Vol. 10560. SPIE. 2018, pp. 189–198.

[80] J. Farina, D. Khomchenko, D. Yevseyenko, J. Meester, and A. Richter. "Automated and comprehensive link engineering supporting branched, ring, and mesh network topologies." In: *Optical Metro Networks and Short-Haul Systems VIII*. Vol. 9773. SPIE. 2016, pp. 134–144.

[81] R. Freund, L. Molle, F. Raub, C. Caspar, M. Karkri, et al. "Triple-(S/C/L)-band WDM transmission using erbium-doped fibre amplifiers." In: *2005 31st European Conference on Optical Communication, ECOC 2005*. Vol. 1. 2005, 69–70 vol.1. DOI: 10.1049/cp:20050382.

[82] A. H. Gnauck, G. Charlet, P. Tran, P. J. Winzer, C. R. Doerr, et al. “25.6-Tb/s C+L-Band Transmission of Polarization-Multiplexed RZ-DQPSK Signals.” In: *Optical Fiber Communication Conference and Exposition and The National Fiber Optic Engineers Conference*. Optica Publishing Group, 2007, PDP19.

[83] A. Sano, H. Masuda, T. Kobayashi, M. Fujiwara, K. Horikoshi, et al. “Ultra-High Capacity WDM Transmission Using Spectrally-Efficient PDM 16-QAM Modulation and C- and Extended L-Band Wideband Optical Amplification.” In: *Journal of Lightwave Technology* 29.4 (2011), pp. 578–586. DOI: 10.1109/JLT.2011.2107030.

[84] R. Sadeghi, B. Correia, E. Virgillito, A. Napoli, N. Costa, et al. “Cost-Effective Capacity Increase of Deployed Optical Networks to Support the Future Internet: the Multi-Band Approach.” In: *2021 12th International Conference on Network of the Future (NoF)*. 2021, pp. 1–7. DOI: 10.1109/NoF52522.2021.9609888.

[85] A. Varasteh, S. K. Patri, A. Autenrieth, and C. Mas-Machuca. “Evaluation of Lightpath Deployment Strategies in Flexible-Grid Optical Networks.” In: *2021 International Conference on Optical Network Design and Modeling (ONDM)*. 2021, pp. 1–6. DOI: 10.23919/ONDM51796.2021.9492493.

[86] D. Uzunidis, E. Kosmatos, C. Matrakidis, A. Stavdas, and A. Lord. “Strategies for Upgrading an Operator’s Backbone Network Beyond the C-Band: Towards Multi-Band Optical Networks.” In: *IEEE Photonics Journal* 13.2 (2021), pp. 1–18. DOI: 10.1109/JPHOT.2021.3054849.

[87] A. Varasteh, S. K. Patri, A. Autenrieth, and C. Mas-Machuca. “Towards Dynamic Network Reconfigurations for Flexible Optical Network Planning.” In: *2021 Optical Fiber Communications Conference and Exhibition (OFC)*. IEEE. 2021, pp. 1–3.

[88] T. Zami, B. Lavigne, and O. B. Pardo. “Benchmarking of Opaque Versus Transparent Core WDM Networks Featuring 400ZR+ QSFP-DD or CFP2 Interfaces.” In: *2020 European Conference on Optical Communications (ECOC)*. 2020, pp. 1–4. DOI: 10.1109/ECOC48923.2020.9333414.

[89] T. Zami and B. Lavigne. “Optimal deployments of 400 Gb/s multihaul CFP2-DCO transponders in transparent IPoWDM core networks.” In: *Optical Fiber Communication Conference (OFC) 2022*. Optica Publishing Group, 2022, Th1F.1. DOI: 10.1364/OFC.2022.Th1F.1.

[90] N. Sambo, P. Castoldi, A. D’Errico, E. Riccardi, A. Pagano, et al. “Next generation sliceable bandwidth variable transponders.” In: *IEEE Communications Magazine* 53.2 (2015), pp. 163–171. DOI: 10.1109/MCOM.2015.7045405.

[91] R. Eisenach. *Understanding 400zr/openzr+/400zr+ Optics*. Accessed: July 2022. URL: https://www.lightwaveonline.com/optical-tech/transmission/article/14188934/understanding-400zropenzr400zr-optics.

[92] S. Melle. *Building IP-optical solutions that are more than the sum of their parts*. https://nokia.ly/3dR3d4D, Accessed: April 2022.

[93] M. Jinno, H. Takara, Y. Sone, K. Yonenaga, and A. Hirano. "Multiflow optical transponder for efficient multilayer optical networking." In: *IEEE Communications Magazine* 50.5 (2012), pp. 56–65. DOI: 10.1109/MCOM.2012.6194383.

[94] J. Zhang, Y. Ji, M. Song, Y. Zhao, X. Yu, et al. "Dynamic traffic grooming in sliceable bandwidth-variable transponder-enabled elastic optical networks." In: *Journal of Lightwave Technology* 33.1 (2015), pp. 183–191.

[95] A. Souza, R. Sadeghi, B. Correia, N. Costa, A. Napoli, et al. "Optimal Pay-As-You-Grow Deployment on S+C+L Multi-band Systems." In: *2022 Optical Fiber Communications Conference and Exhibition (OFC)*. 2022, pp. 1–3.

[96] J. Alberto Hernandez, M. Quagliotti, L. Serra, L. Luque, R. Lopez da Silva, et al. "Comprehensive model for technoeconomic studies of next-generation central offices for metro networks." In: *Journal of Optical Communications and Networking* 12.12 (2020), pp. 414–427. DOI: 10.1364/JOCN.402167.

[97] M. Cantono, R. Schmogrow, M. Newland, V. Vusirikala, and T. Hofmeister. "Opportunities and challenges of C+ L transmission systems." In: *Journal of Lightwave Technology* 38.5 (2020), pp. 1050–1060.

[98] K. Jordan. *The benefits of an integrated C&L-band photonic line system*. https://www.ciena.com/insights/articles/The-benefits-of-an-integrated-CL-band-photonic-line-system.html Accessed: August 2022.

[99] R. K. Jana, A. Mitra, A. Pradhan, K. Grattan, A. Srivastava, et al. "When Is Operation Over C+L Bands More Economical than Multifiber for Capacity Upgrade of an Optical Backbone Network?" In: *2020 European Conference on Optical Communications (ECOC)*. 2020, pp. 1–4. DOI: 10.1109/ECOC48923.2020.9333276.

[100] I. Busi. *IP + Optical: The Mainstream Solution for the 400G Era*. https://blog.huawei.com/2022/07/04/ip-optical-mainstream-solution-400g/, Accessed: August 2022.

[101] P. Morales, P. Franco, A. Lozada, N. Jara, F. Calderón, et al. "Multi-band Environments for Optical Reinforcement Learning Gym for Resource Allocation in Elastic Optical Networks." In: *2021 International Conference on Optical Network Design and Modeling (ONDM)*. 2021, pp. 1–6. DOI: 10.23919/ONDM51796.2021.9492435.

[102] M. Nakagawa, T. Seki, and T. Miyamura. “Techno-Economic Potential of Wavelength-Selective Band-Switchable OXC in S+C+L Band Optical Networks.” In: *2022 Optical Fiber Communications Conference and Exhibition (OFC)*. 2022, pp. 01–03.

[103] J. M. Simmons and G. N. Rouskas. “Routing and Wavelength (Spectrum) Assignment.” In: *Springer Handbook of Optical Networks*. Springer, 2020, pp. 447–484.

[104] H. Lun, M. Fu, X. Liu, Y. Wu, L. Yi, et al. “Soft failure identification for long-haul optical communication systems based on onedimensional convolutional neural network.” en. In: *IEEE/OSA Journal of Lightwave Technology* 38.11 (2020). LFL+20, pp. 2992–2999.

[105] S. Shahkarami, F. Musumeci, F. Cugini, and M. Tornatore. “Machinelearning- based soft-failure detection and identification in optical networks.” en. In: *2018 Optical Fiber Communications Conference and Exposition (OFC*. 2018, pp. 1–3.

[106] K. Grobe, C. Furst, A. Autenrieth, and T. Szyrkowiec. “Flexible spectrum-as-a-service.” en. In: *TERENA Networking Conference (TNC*. May 2014.

[107] K. Kaeval, T. Fehenberger, J. Zou, S. L. Jansen, K. Grobe, et al. “QoT assessment of the optical spectrum as a service in disaggregated network scenarios.” In: *J. Opt. Commun. Netw.* 13.10 (Oct. 2021), E1–E12. DOI: 10.1364/JOCN.423530. URL: https://opg.optica.org/jocn/abstract.cfm?URI=jocn-13-10-E1.

[108] Z. Wang, M. Zhang, D. Wang, C. Song, M. Liu, et al. “Failure prediction using machine learning and time series in optical network.” en. In: *Optics Express* 25.16 (Aug. 2017), pp. 18553–18565.

[109] C. Natalino, A. Udalcovs, L. Wosinska, O. Ozolins, and M. Furdek. “Spectrum anomaly detection for optical network monitoring using deep unsupervised learning.” en. In: *IEEE Communications Letters* (2021). NUW+21, pp. 1–1.

[110] D. Rafique, T. Szyrkowiec, A. Autenrieth, and J. Elbers. “Analytics-driven fault discovery and diagnosis for cognitive root cause analysis.” en. In: *2018 Optical Fiber Communications Conference and Exposition (OFC*. RSAE18. 2018, pp. 1–3.

[111] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, et al. “Scikit-learn: Machine learning in python.” en. In: *Journal of Machine Learning Research* 12.85 (2011), pp. 2825–2830.

[112] R. Hyndman and G. Athanasopoulos. “Forecasting: principles and practice.” es. In: *OTexts* (2018).

[113] J. deLeeuw. "Introduction to Akaike (1973) information theory and an extension of the maximum likelihood principle." In: *Breakthroughs in statistics*. Springer, 1992, pp. 599–609.

[114] J. Bergstra and Y. Bengio. "Random search for hyper-parameter optimization." en. In: *The Journal of Machine Learning Research* 13.281-305 (2012), p. 03.

[115] *PR Newswire, The Worldwide Optical Communication and Networking Equipment Industry is Expected to Reach $36.6 Billion by 2027*. https://prn.to/3S7z8xt, Accessed: July 2022.

[116] S. Melin et al. *TIP OOPT MUST Optical Whitepaper*. https://tinyurl.com/y63n4hce, Accessed: July 2022.

[117] A. Triki et al. "OpenROADM Compliant SDN Controller for a Full Interoperability of the Optical Transport Network." In: *2018 European Conference on Optical Communication (ECOC)*. 2018, pp. 1–3. DOI: 10.1109/ECOC.2018.8535561.

[118] M. Svaluto Moreolo, J. M. Fabrega, L. Nadal, F. J. Vílchez, A. Mayoral, et al. "SDN-Enabled Sliceable BVT Based on Multicarrier Technology for Multiflow Rate Distance and Grid Adaptation." In: *Journal of Lightwave Technology* 34.6 (2016), pp. 1516–1522. DOI: 10.1109/JLT.2015.2510082.

[119] S. Troia et al. "Machine-learning-assisted Routing in SDN-based Optical Networks." In: *2018 European Conference on Optical Communication (ECOC)*. IEEE. 2018, pp. 1–3.

[120] R. Vilalta et al. "Architecture to Deploy and Operate a Digital Twin Optical Network." In: *Optical Fiber Communication Conference*. Optica Publishing Group. 2022, W1F–4.

[121] OpenConfig. *OpenConfig*. https://www.openconfig.net/. Accessed: July 2022.

[122] R. Vilalta, R. Casellas, R. Martínez, R. Muñoz, A. Gonzalez-Muñiz, et al. "Optical Network Telemetry with Streaming Mechanisms using Transport API and Kafka." In: *2021 European Conference on Optical Communication (ECOC)*. 2021, pp. 1–4. DOI: 10.1109/ECOC52684.2021.9606002.

[123] F. Cugini, F. Paolucci, G. Meloni, G. Berrettini, M. Secondini, et al. "Push-Pull Defragmentation Without Traffic Disruption in Flexible Grid Optical Networks." In: *Journal of Lightwave Technology* 31.1 (2013), pp. 125–133. DOI: 10.1109/JLT.2012.2225600.

[124] *RFC 8526 - NETCONF Extensions to Support the Network Management Datastore Architecture*. Tech. rep. (Accessed on 02/27/2023).

[125] *RFC 8040 - RESTCONF Protocol*. Tech. rep. (Accessed on 02/27/2023).

[126] highstreet-technologies GmbH. *highstreet-technologies TransportPCE*. https://github.com/highstreet-technologies/transportpce/tree/opticon, Accessed: July 2022.

[127] *Kevin Mitnick demonstrates how easy it is for a hacker read your email messages - YouTube*. https://www.youtube.com/watch?v=KcJWXpABpVo&ab_channel=IsaacP.. (Accessed on 02/23/2023).

[128] V. Bhatia and K. Ramkumar. "An Efficient Quantum Computing technique for cracking RSA using Shor's Algorithm." In: *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*. 2020. DOI: 10.1109/ICCCA49541.2020.9250806.

[129] M. Gunkel, F. Wissel, and A. Poppe. "Designing a Quantum Key Distribution Network-Methodology and Challenges." In: *Photonic Networks; 20th ITG-Symposium*. VDE. 2019, pp. 1–3.

[130] Y. Cao, Y. Zhao, J. Wang, X. Yu, Z. Ma, et al. "Cost-efficient quantum key distribution (QKD) over WDM networks." In: *Journal of Optical Communications and Networking* 11.6 (2019), pp. 285–298.

[131] Y. Cao, Y. Zhao, Q. Wang, J. Zhang, S. X. Ng, et al. "The evolution of quantum key distribution networks: On the road to the qinternet." In: *IEEE Communications Surveys & Tutorials* 24.2 (2022), pp. 839–894.

[132] F. Pederzolli, F. Faticanti, and D. Siracusa. "Optimal Design of Practical Quantum Key Distribution Backbones for Securing CoreTransport Networks." In: *Quantum Reports* 2 (Jan. 2020), pp. 114–125. DOI: 10.3390/quantum2010009.

[133] Q. Li, Y. Wang, H. Mao, J. Yao, and Q. Han. "Mathematical model and topology evaluation of quantum key distribution network." In: *Opt. Express* 28.7 (Mar. 2020), pp. 9419–9434. DOI: 10.1364/OE.387697. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-28-7-9419.

[134] B. Karanov, T. Xu, N. A. Shevchenko, D. Lavery, R. I. Killey, et al. "Span length and information rate optimisation in optical transmission systems using single-channel digital backpropagation." In: *Optics Express* 25.21 (2017), pp. 25353–25362.

[135] Y. Ou, E. Hugues-Salas, F. Ntavou, R. Wang, Y. Bi, et al. "Field-Trial of Machine Learning-Assisted Quantum Key Distribution (QKD) Networking with SDN." In: *2018 European Conference on Optical Communication (ECOC)*. 2018, pp. 1–3. DOI: 10.1109/ECOC.2018.8535497.

[136] E. Crockett, C. Paquin, and D. Stebila. *Prototyping post-quantum and hybrid key exchange and authentication in TLS and SSH*. Cryptology ePrint Archive, Paper 2019/858. https://eprint.iacr.org/2019/858. 2019. URL: https://eprint.iacr.org/2019/858.

[137] C. Paquin, D. Stebila, and G. Tamvada. "Benchmarking post-quantum cryptography in TLS." In: *Proc. 11th International Conference on Post-Quantum Cryptography (PQCrypto) 2020*. Ed. by J. Ding and J.-P. Tillich. Vol. 12100. LNCS. Springer, 2020, pp. 72–91.

[138] D. Stebila and M. Mosca. "Post-quantum key exchange for the Internet and the Open Quantum Safe project." In: *Proc. 23rd Conference on Selected Areas in Cryptography (SAC) 2016*. Ed. by R. Avanzi and H. Heys. Vol. 10532. LNCS. Springer, 2017, pp. 1–24. DOI: 10.1007/978-3-319-69453-52.

[139] G. Alagic, D. Apon, D. Cooper, Q. Dang, T. Dang, et al. "Status report on the third round of the NIST post-quantum cryptography standardization process." In: *US Department of Commerce, NIST* (2022).

[140] D. A. Cooper, D. C. Apon, Q. H. Dang, M. S. Davidson, M. J. Dworkin, et al. "Recommendation for stateful hash-based signature schemes." In: *NIST Special Publication* 800 (2020), p. 208.

[141] J. Buchmann, E. Dahmen, and A. Hülsing. "XMSS-a practical forward secure signature scheme based on minimal security assumptions." In: *Post-Quantum Cryptography: 4th International Workshop, PQCrypto 2011, Taipei, Taiwan, November 29–December 2, 2011. Proceedings 4*. Springer. 2011, pp. 117–129.

[142] D. McGrew, M. Curcio, and S. Fluhrer. *Leighton-Micali hash-based signatures*. Tech. rep. 2019.

[143] A. Shamir. "How to share a secret." In: *Communications of the ACM* 22.11 (1979), pp. 612–613.

[144] A. V. Paz. "Link Migration Strategies for future-ready Optical Transport Networks." M.Sc. Thesis. Munich, Germany: Chair of Communication Networks, Technical University of Munich, 2023.

[145] I. F. Dick. "Machine Learning-based Fault Detection Algorithms for Long Haul Elastic Optical Networks." M.Sc. Thesis. Munich, Germany: Chair of Communication Networks, Technical University of Munich, 2021.

[146] E. Bilgin. *Mastering Reinforcement Learning with Python*. 2020. ISBN: 978-1-83864-414-7. URL: https://learning.oreilly.com/library/view/mastering-reinforcement-learning/9781838644147/ (Last Accessed: 2.9.2022).

[147] P. Winder. *Reinforcement Learning*. 2020. ISBN: 978-1-4920-7238-6. URL: https://learning.oreilly.com/library/view/reinforcement-learning/9781492072386/ (Last Accessed: 2.9.2022).

[148] S. Mustafa. *Reinforcement Learning and Deep RL Python (Theory and Projects)*. 2022. URL: https://learning.oreilly.com/videos/reinforcement-learning-and/9781804610626/ (Last Accessed: 27.2.2023).

[149] S. Ravichandiran. *Deep Reinforcement Learning with Python - Second Edition*. 2020. ISBN: 978-1-83921-068-6. URL: https://learning.oreilly.com/library/view/deep-reinforcement-learning/9781839210686/ (Last Accessed: 27.2.2023).

List of Figures

List of Tables