8-11-2015

Management of Spectral Resources in Elastic Optical Networks

Sunny Shakya

TITLE: MANAGEMENT OF SPECTRAL RESOURCES IN ELASTIC OPTICAL NETWORKS

by

SUNNY SHAKYA

Under the Direction of Xiaojun Cao, PhD

ABSTRACT

Recent developments in the area of mobile technologies, data center networks, cloud computing and social networks have triggered the growth of a wide range of network applications. The data rate of these applications also vary from a few megabits per second (Mbps) to several Gigabits per second (Gbps), thereby increasing the burden on the Internet. To support this growth in Internet data traffic, one foremost solution is to utilize the advancements in optical networks. With technology such as wavelength division multiplex-

ing (WDM) networks, bandwidth upto 100 Gbps can be exploited from the optical fiber in an energy efficient manner. However, WDM networks are not efficient when the traffic demands vary frequently. Elastic Optical Networks (EONs) or Spectrum Sliced Elastic Optical Path Networks (SLICE) or Flex-Grid has been recently proposed as a long-term solution to handle the ever-increasing data traffic and the diverse demand range. EONs provide abundant bandwidth by managing the spectrum resources as fine-granular orthogonal sub-carriers that makes it suitable to accommodate varying traffic demands. However, the Routing and Spectrum Allocation (RSA) algorithm in EONs has to follow additional constraints while allocating sub-carriers to demands. These constraints increase the complexity of RSA in EONs and also, make EONs prone to the fragmentation of spectral resources, thereby decreasing the spectral efficiency. The major objective of this dissertation is to study the problem of spectrum allocation in EONs under various network conditions. With this objective, this dissertation presents the author's study and research on multiple aspects of spectrum allocation in EONs: how to allocate sub-carriers to the traffic demands, how to accommodate traffic demands that varies with time, how to minimize the fragmentation of spectral resources and how to efficiently integrate the predictability of user demands for spectrum assignment. Another important contribution of this dissertation is the application of EONs as one of the substrate technologies for network virtualization.

INDEX WORDS: Elastic Optical Networks, Flex-Grid networks, Spectrum-sliced elastic optical networks, Routing and Spectrum Allocation, Spectrum fragmentation, Network virtualization, Virtual network reconfiguration

TITLE: MANAGEMENT OF SPECTRAL RESOURCES IN ELASTIC OPTICAL NETWORKS

by

SUNNY SHAKYA

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2015

TITLE: MANAGEMENT OF SPECTRAL RESOURCES IN ELASTIC OPTICAL NETWORKS

by

SUNNY SHAKYA

Committee Chair: Xiaojun Cao

Committee: Raj Sunderraman

Yanqing Zhang

Hendricus Van Der Holst

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

August 2015

DEDICATION

I want to dedicate this dissertation to my wife, Mrs. Nabina Pradhan Shakya for encouraging me to pursue this PhD degree and for all her love and support during this period, and to my parents and brother for their unconditional love.

ACKNOWLEDGMENTS

This dissertation work would not have been possible without the support and guidance of many people. I want to express my sincere gratitude towards my advisor, Dr. Xiaojun Cao for providing me with a research direction and for his time, guidance and expertise throughout this dissertation. His constant guidance and suggestions on writing paper, giving talks and presentation and solving complex mathematical equations have helped me get this far. I would also like to thank Dr. Raj Sunderraman for all his encouragement and suggestions. The excitement he has for his work has given me a lot of energy to carry on this dissertation. Lastly, I would like to thank my PhD committee members for providing me with valuable suggestions on the overall form of this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

- SLICE - Spectrum-sliced Elastic Optical Networks
- EONs - Elastic Optical Networks
- WDM - Wavelength Division Multiplexing
- WRN - Wavelength Routed Network
- EDFA - Erbium Doped fiber amplifier
- ADM - Add/Drop Multiplexers
- WB -Waveband Switching
- DWDM - Dense Wavelength Division Multiplexing
- WXCs - Wavelength Cross Connects
- OPS - Optical Packet Switching
- OBS - Optical Burst Switching
- OFDM - Orthogonal Frequency Division Multiplexing
- RWA - Routing and Wavelength Assignment
- RSA - Routing and Spectrum Assignment
- SPs - Spectrum Paths
- MIS - Maximum Independent Sets
- MSI - Maximum-used Sub-carrier Index
- IG - Interference Graph

- GB - Guard Bands
- SRA - Static Routing Assignment
- ILP - Integer Linear Programming
- BPSK - Binary Phase Shift Keying
- QPSK - Quadrature Phase Shift Keying
- QAM - Quadrature Amplitude Modulation
- MAMR - Modulation-aware Multipath Routing Algorithm
- TSA - Time Varying Spectrum Allocation
- SLE - Static Lightpath Assignment
- SAS - Spectrum Allocation from Scratch
- PSA - Predictive Spectrum Allocation
- PPSA - Plan-ahead Predictive Spectrum Allocation
- PSBA - Predictive Spectrum Band Allocation
- VN - Virtual Network
- VNE - Virtual Network Embedding
- ACT-VNE - Alignment and Consecutiveness-aware Virtual Network Embedding
- iACT-VNE - Importance, Alignment and Consecutiveness-aware Virtual Network Embedding
- VNR - Virtual Network Reconfiguration
- RCLM-VNR - Relative Consecutiveness Loss-aware and Misalignment-aware Virtual Network Reconfiguration

PART 1

INTRODUCTION

The Internet data traffic has been growing at a tremendous rate since its inception. The recent advancements in the areas of mobile technologies, data center networks and cloud computing have brought a huge amount of emerging applications along with a huge amount of data traffic to Internet. Cisco estimates that the global Internet traffic has increased more than five folds in the past 5 years, and predicts that the Internet traffic will increase threefold over the next 5 years [2]. This exponential growth in Internet traffic poses a great threat to the underlying network infrastructure and calls for a more scalable, agile and robust infrastructure to support the future Internet data traffic.

Optical networks that use optical fibers to carry the data in the form of light have been considered a promising solution to handle the Internet traffic growth as it offers a huge amount of bandwidth. With the state-of-the-art wavelength division multiplexing (WDM) technology, each fiber can carry wavelengths with bandwidth up to 100 Gbits/s or higher [3], [4]. To satisfy the user demands, wavelengths are assigned and routed to form lightpaths via the Routing and Wavelength assignment (RWA) algorithms, which result in the wavelength routed network (WRN). Despite the popularity of WDM networks, the coarse and rigid wavelength assignment in WDM networks is considered to be spectrally inefficient [1], [5]. WDM networks divide the fiber frequencies into channels with a fixed-sized spectral bandwidth. This rigid and coarse resource allocation leads to a poor utilization of the spectral resources when the traffic demand between any two optical nodes varies significantly. Hence, the sub-wavelength traffic may have to be over-provisioned due to the coarse granularity of the wavelength. Besides this, when a traffic demand requires multiple wavelengths (i.e., super-wavelength traffic), guard-band frequencies between two wavelengths may lead to the under-utilization of the available spectrum resources as shown

in Fig. 1.1. To make more efficient and elastic use of the spectrum, Spectrum-sLICed Elastic optical path (SLICE) networks or Elastic Optical Networks (EONs), or Flex-Grid networks has been proposed recently.

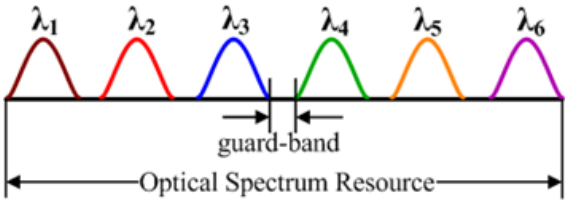


Figure (1.1) An example of wavelength assignment in WDM networks

EONs employ the orthogonal frequency division multiplexing (OFDM) technology and divide the spectrum resources into fine-granular spectrum bands known as sub-carriers for better utilization of spectral resources [1], [5], [6], [7]. Unlike the WDM networks, EONs carry data on a number of sub-carriers that are overlapped in the frequency domain because of their orthogonality property as shown in Figure 1.2. Each sub-carrier has a much lower granularity than a single wavelength which makes it one of the better candidates to optically accommodate the sub-wavelength traffic [6]. The super-wavelength traffic can be accommodated by combining several sub-carriers together without placing the guard frequencies (or carriers) in between. EONs increase the flexibility while assigning sub-carriers to sub-wavelength and super-wavelength traffic. Furthermore, in EONs, a lightpath can expand or contract elastically to meet the needs of varying traffic demands. However, EONs pose

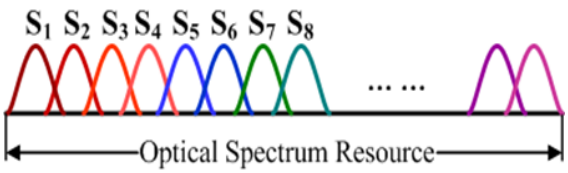


Figure (1.2) An example of sub-carrier distribution in EONs

additional constraints to be satisfied on the networking level, specifically on the efficient establishment of connections and the assignment of sub-carriers to the traffic demands.

The routing and spectrum allocation (RSA) algorithm in EONs allocates a spectrum path (SP) to the incoming connection demand similar to the Routing and Wavelength Allocation (RWA) algorithm in WRN. The RSA algorithm must ensure that there is a continuous availability of sub-carriers along a SP (if no frequency converter is provided), which is called *spectrum-continuity* constraint. In addition, the sub-carriers assigned to any SP should be consecutive in spectrum domain to take full advantages of the OFDM technology. This new contiguity constraint known as the *sub-carrier consecutiveness* constraint adds a degree of complexity to the conventional RWA problem. Further, the physical frequency filtering requires that various SPs in EONs are separated in the spectrum domain by guard frequencies when two or more SPs share one or more common fiber links. These guard frequencies are referred to as the guard-carriers. However, unlike the WDM network, where the guard-band frequencies are pre-allocated and fixed, the guard-carriers in EONs can be any of the sub-carriers and are determined in the process of spectrum paths establishment. The existence of the *sub-carrier consecutiveness* and *guard-carrier* constraint further complicate the process of RSA in EONs.

In EONs with dynamic traffic, the frequent set-up and tear down of SPs may lead to significant fragmentation of spectral resources [8], [9]. Due to the *spectrum-continuity* and *sub-carrier consecutiveness* constraints, several spectrum slots in between connections remain unused. These spectral fragments are small, scattered and may not be enough to establish new connections and decrease the probability of finding sufficient contiguous sub-carriers for new connections. New arrival of requests are then forced either to utilize more spectrum in the fiber or are blocked even though sufficient spectrum are available. Hence, the spectral fragments represent the wastage of the expensive spectral resources which may lower the spectrum utilization and also, increase the blocking probability. In the light of this, there is an increasing demand from network operators to be able to periodically reconfigure their networks, aiming to improve the spectrum utilization. This operation is called spectrum

defragmentation. During the process of spectrum defragmentation, the available fragmented spectrum bands are consolidated by reconfiguring the active connections, i.e., changing their routes, or assigning them a different set of sub-carriers or both, while maintaining the imposed continuity and contiguity constraints. The presence of spectrum fragmentation in EONs further complicates the process of RSA in EONs.

This dissertation mainly studies the problem of the management of spectral resources in EONs under various network conditions. Using this study, this dissertation proposes novel RSA algorithms in EONs that use the predictability in Internet data traffic to assign sub-carriers to demands with minimum interruptions to the already existing connections. Owning to the fact that EONs are prone to spectrum fragmentation, this dissertation also proposes novel spectrum defragmentation algorithms to minimize spectrum fragmentation. Since spectrum fragmentation is inevitable in EONs, this dissertation also presents a novel multipath routing algorithm that can allocate sub-carriers to demands along multiple routing paths in order to minimize the blocking probability. Finally, this dissertation proposes novel algorithms that aid in the application of EONs in network virtualization. The application of EONs in network virtualization enables the agile and scalable management of networks as a service allowing multiple heterogeneous networks to co-exist together in isolation with each other.

The rest of this dissertation is organized as follows. Chapter II briefly explains the evolution of EONs and also, discusses some of the enabling technologies and the underlying architecture for EONs. Chapter III introduces the problem of RSA in EONs and also, explains various constraints and conditions RSA has to fulfill and briefly goes through various RSA algorithms proposed in the literature. Thereafter, each individual chapter deals with a specific issue. Along this process, the problem under study is stated, followed by a review of related existing work in the literature. Next, contributions addressing the identified issue are provided and further validated by numerical experiments on various network models. In this regard, Chapter IV presents the problem of spectral fragmentation in EONs and proposes novel multipath algorithms that can cope with fragmented resources and also,

propose spectral defragmentation algorithm. Chapter V presents novel sub-carrier allocation algorithms to allocate spectral resources to fluctuating demands based on the traffic condition and predictability of Internet traffic demand. Chapter VI presents the recent research in the application of EONs in network virtualization and proposes various strategies for this application. Finally, the main highlight and the impact of the overall study are concluded in Chapter VII.

PART 2

THE EVOLUTION OF ELASTIC OPTICAL NETWORKS

In this chapter, we present the detailed overview on both the history and the evolution of EONs. In this chapter, we take a detailed look at why we need a more scalable and agile technology for carrying the data in optical networks even though WDM technology has been already applied in optical networks for more than a decade. We also take a look at some of the remarkable inventions and the underlying technologies that leads towards the introduction of EONs and why EONs are better than WDM networks.

2.1 Early efforts in WDM and optical switching

The recent developments in the area of mobile technologies, social networks, data center networks and cloud computing have increased the amount of Internet data traffic by many folds. With the forecast of more explosive growth in Internet data traffic in future years [2], a transmission medium that can handle the current data traffic and also, support the future traffic growth is mandatory. Optical communication systems with optical fiber as a transmission medium offers a very promising solution to handle the Internet data traffic growth as optical fibers have huge amount of bandwidth to offer. The optical networks have also come a long way and have gone through a rapid evolution since the first optical transmission system was commercially demonstrated in 1980 at a line rate of 45 Mb/s [10], [11], [12]. It was a single wavelength system made up of a series of optical links - a transmitter, a fiber line, and a receiver. After several generations and with the innovations of many important devices, the operation of optical transmission systems at 12 Gb/s became possible. WDM networks was proposed very early in the evolution of optical transmission systems. Under WDM, the available optical spectrum in the optical fiber is divided into a number of non-overlapping wavelengths, with each wavelength supporting a single communication

channel. However, due to the lack of enabling technologies for WDM networks such as optical amplifiers and electrical regenerators, WDM networks remained a distant technology. Early researchers demonstrated the operation of basic optical switches, both bulk and integrated. These optical switches could provide a wavelength add/drop functionality. More complex version of optical switches such as the Optical Cross Connect (OXC) switch fabric was also developed that used a switching matrix to switch a particular wavelength from one of N input fiber routes to any of N output fiber routes (e.g. Reconfigurable Optical Add/Drop Multiplexers (ROADM) [13]). Later, the invention of optical amplifiers aided the operation of WDM networks. The earliest and most successful optical fiber amplifier is the Erbium Doped Fiber Amplifiers (EDFA). The optical gain achieved with EDFA and is sufficient to compensate for loss of a 100-km span which made the EDFA an ideal amplifier for high capacity WDM transmission systems. While optical transmission mostly benefited from the availability of optical amplifier, complementary optical switching capability was largely absent from the first-generation network nodes.

Commercial deployment of the point-to-point WDM transmission systems came in 1995, which included only eight channels at 2.5 Gb/s. However, to stay ahead of perceived demand, 16 wavelength systems with a total capacity of 40 Gb/s were deployed in 1997. With WDM networks, several optical signals can be transmitted on the same optical fiber using different wavelength channels which further expand the capacity of optical fibers. Initially in optical communications, the WDM technology was used to increase the capacity of point-to-point optical fibers/links where the signal is converted back to the electrical domain at the end of each fiber/link. Since this conversion is slow and expensive, the trend has been towards transparent all-optical networks, where the signal is routed in the optical domain through-out the entire network. The invention of the optical fiber amplifier (EDFA) made this possible. The optical fiber amplifier is capable of amplifying several optical signals at the same time without undergoing conversion into the electrical domain. Further, a wavelength-routed network (WRN) was realized with the help of Add/Drop Multiplexers (ADM). A wavelength-routed network (WRN) is an all-optical network, where the routing at the network nodes is

based on the wavelength of the incoming signal [3], [4].

The introduction of polarization division multiplexing, phase and multi-level modulation formats, coherent detection and digital equalization in electrical domain along with the advancement in optical amplification enabled long-distance dense wavelength-division multiplexed (DWDM) transmission with per-channel bandwidth of 100 Gb/s [1]. The extended distance, an optical signal can travel through multiple DWDM links and wavelength cross-connects (WXCs) without undergoing optical-electrical optical (OEO) regeneration has made the optically routed networks feasible. The optically routed networks have some major advantages such as the elimination of OEO regenerators that are costly, space and power-consuming and automated remote provisioning of optical paths. However, the coarse allocation and the large granularity of the wavelength was major concerns in the realization of optical communication systems. The WRN requires full allocation of an entire wavelength to an optical path even when the traffic demand is not sufficient to occupy the entire capacity of wavelength (sub-wavelength traffic). Further, when the demand is greater than the capacity of a single wavelength (super-wavelength traffic), several wavelengths can be grouped together in a scalable manner. However, adjacent wavelengths have to be separated by guardbands (GB), which can be in the order of several wavelengths. This may result in wastage of the expensive spectral resources. This led to the development of several other approaches such as optical packet switching (OPS) [14], [15], optical burst switching (OBS) [16] and waveband switching (WS) [17] to meet the requirements of sub-wavelength and super-wavelength traffic accommodation.

Waveband switching [17] can be applied to reduce the port count of optical cross-connects (OXC) by switching a group of wavelengths together as a waveband using a single cross-connect port. Accordingly, a non-uniform waveband was proposed for an efficient accommodation of a wide range of traffic [18]. This significantly reduces the count of the port number in OXC. However, wavelength demultiplexing requires that the adjacent wavelengths have to be separated by a buffer in the spectral domain. This leads to low spectral efficiency. In OPS networks, all-optical data packets that are composed of an optical header

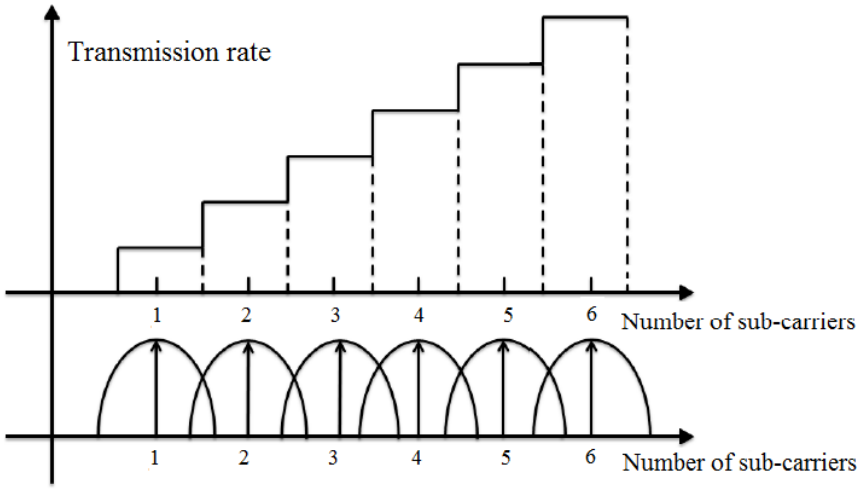


Figure (2.1) Variable bandwidth transmission with OFDM sub-carriers.

and an optical payload are statistically-multiplexed in the DWDM optical layer, so that the wavelength in each link is shared amongst all packets belonging to all source-destination demand pairs. Nevertheless, OPS cannot be realized due to the lack of several key technological [14]. First, no optical random access memories (RAM) exists to date, which restricts packet store-and-forward operations as in IP routers. The only solution is to delay packets in the optical domain by using fiber delay lines (FDLs). However, these devices are bulky and only provide deterministic and minimal delays. In addition, optical header processing is also not feasible today. All these technological shortcomings prevented the realization of OPS. To relax the need for optical RAM, OBS network has been proposed as a suitable candidate for next-generation all-optical networks. In OBS [16], several small bursts are combined into one big burst at the ingress router and routed towards the egress router using only one control packet. This leads to a more data per header than in packet networks. Therefore, a higher data rate is obtained with the same header processing rate. Additionally, no complex buffering elements are necessarily needed as bursts can be buffered in electronics form at the edge of the network instead of buffering at each intermediate node. However, approaches such as OBS and OPS that meet these requirements can only be viewed as long-term solutions since the enabling technologies for these approaches are still immature.

2.2 Elastic optical network: architecture, benefits, and enabling technologies

Orthogonal Frequency-Division Multiplexing (OFDM) has been recently proposed as a modulation technique in optical networks [1]. In OFDM, the data is transmitted over multiple orthogonal sub-carriers with finer granularity. Besides the advantages of low symbol rate of each sub-carrier, OFDM also brings unique benefits in terms of spectral efficiency, allowing adjacent sub-carriers to overlap each other due to their orthogonal modulation. This unique feature in OFDM enables the elastic bandwidth transmission by allocating a variable number of low-rate sub-carriers for a transmission as shown in Figure 2.1.

OFDM together with bandwidth-variable (BV) transponders and bandwidth variable-wavelength cross-connects (BV-WXC) paved the way for the elastic allocation of spectral resources. To form a connection between two optical nodes, BV transponders may be deployed at the network edges whereas BC-WXC are deployed at the network core. BV transponder generates an optical signal using just enough spectral resources, in terms of sub-carriers with appropriate modulation level to satisfy the traffic demand. To establish a connection, every BV-WXC on the route allocates a cross-connection with sufficient spectrum to create an appropriately sized end-to-end optical path. For this, the BV-WXC [19] has to configure its switching window in a contiguous manner according to the spectral width of the incoming optical signal. Figure 2.2 presents a schematic diagram of BV-WXC.

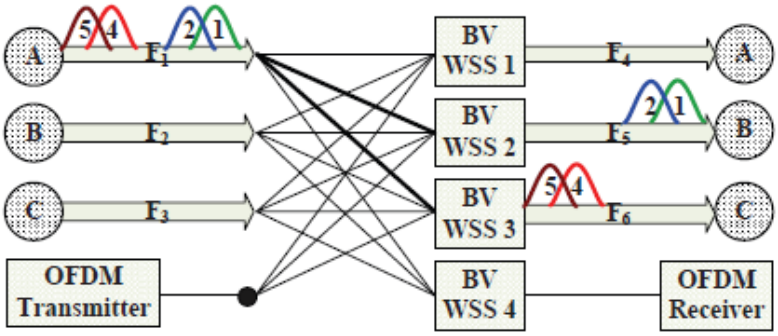


Figure (2.2) Bandwidth variable WXC.

The development of these enabling technologies have enabled a novel spectrum efficient

and scalable optical transport network architecture known as elastic optical networks (EONs) or Spectrum-sliced Elastic Optical Networks (SLICE) or Flex-Grid networks [1]. In EONs, an appropriately-sized optical spectrum is allocated to every connection depending on the traffic demand. Unlike the rigid bandwidth of the conventional fixed-bandwidth optical path in WDM networks, an optical path in EONs can expand and contract according to the traffic demand, if necessary. Such elasticity in optical spectrum allocation significantly improves the spectrum efficiency and scalability of the network. It also provides unique features such as segmentation and aggregation of spectral resources, efficient accommodation of multiple data rates, as well as elastic variation of allocated resources [5]. Figure 2.3 illustrates these features of EONs.

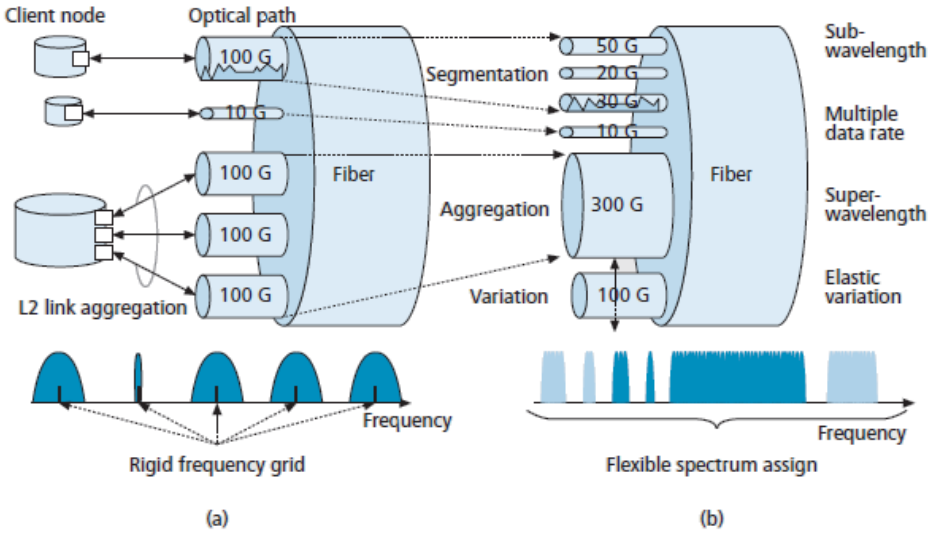


Figure (2.3) Spectrum assignment in (a) WDM networks (b) EONs. [1]

2.2.1 Sub-wavelength traffic accommodation

WDM networks allocate an entire wavelength to a traffic demand even though the demand is not sufficient to fill the entire capacity of the wavelength (sub-wavelength traffic) as shown in Figure 2.3(a). This rigid and coarse assignment of wavelengths leads to an inefficient utilization of spectral resources. EONs allocate just enough sub-carriers to accommodate the traffic demand as shown in Figure 2.3(b). Beside this, the granularity of the sub-carriers

is very small as compared to a wavelength. This makes EONs a suitable candidate to handle the sub-wavelength traffic.

2.2.2 Super-wavelength traffic accommodation

WDM networks allocate multiple independent channels comprising of a number of wavelengths for traffic demands that are greater than a capacity of a wavelength. However, the adjacent wavelengths have to be separated by guard bands (GB) and the size of GB can be in the order of wavelengths. However, in EONs, multiple contiguous sub-carriers can be combined to satisfy the specific demands as shown in Figure 2.3(b). These sub-carriers are overlapping in frequency domain which further increases the spectral efficiency in EONs.

2.2.3 Multiple data rate accommodation

In addition to sub-wavelength and super-wavelength traffic accommodation, EONs also enable spectrally-efficient direct accommodation of mixed data bit rates in the optical domain because of the flexible assignment of spectrum as shown in Figure 2.3. In contrast, WDM networks with fixed grid can lead to stranding of the optical bandwidth due to the excess frequency spacing for lower bit rate signals. In this way, EONs support various data rates including possible future ones in a highly spectrum-efficient manner.

2.3 Chapter Summary

This chapter presented the brief history and the evolution of the enabling technologies that paved the way for the realization of EONs. While EONs introduced a new degree of freedom to the future optical transport networks, it poses additional challenges on the networking level, specifically on the efficient establishment of network routes and spectrum assignment. In the view of this, the RSA problem in EONs is presented in detail in the next chapter.

PART 3

ROUTING AND SPECTRUM ALLOCATION IN ELASTIC OPTICAL NETWORKS

EONs are foreseen as a promising solution for future optical transport. However, it also brings with itself new technical challenges; like the efficient allocation of spectral resources to the elastic lightpaths, contiguous sub-carrier allocation within each lightpaths. Similar to WDM networks, an end-to-end connection in EONs must occupy the same spectrum between its end-nodes, that is, ensuring the *spectrum-continuity* constraint. In addition, the entire spectrum allocated to the connection demands must be contiguously, which is referred to as the *sub-carrier consecutiveness* constraint. This new continuity constraint adds a degree of complexity to the conventional RWA problem. Furthermore, the available RWA algorithms proposed for WDM networks cannot be directly applicable in EONs. A new routing and resource allocation scheme has to be developed, namely RSA. In this chapter, we present the detailed study of the RSA problem in EONs. This chapter, after reviewing the RSA problem in EONs, puts an emphasis on the RSA in static and dynamic network conditions. In this chapter, we also review various heuristic algorithms for RSA and the variation of RSA problem that have been proposed in literature.

3.1 Routing and Spectrum Allocation

In WDM networks, a RWA algorithm assigns a routing path between two optical nodes and assigns the wavelength along that path following the wavelength-continuity constraint. The wavelength-continuity constraint requires that the wavelength assigned to the routing path should be continuous along that path. The resulting network is referred to as the wavelength routed network (WRN) [3], [4]. As already stated, the fixed size allocation of frequency channel in WRN has drawbacks in its coarse granularity and limited flexibility.

In EONs, a frequency channel is divided into finely granular frequency slots (or sub-carriers) which have much less granularity than a single wavelength. Since EONs use OFDM as the modulation format, these sub-carrier can overlap each other in frequency domain and can support a direct accommodation of mixed data bit rates in the optical domain [1], [5]. When assigning a routing path between two optical nodes, the spectrum has to be continuous along the path. This is similar to the wavelength-continuity constraint in WRN. In addition to this, the sub-carriers assigned along that path have to be contiguous in the frequency domain to be modulated effectively. This constraint makes RSA in EONs even more difficult barring RWA proposed for WDM networks to be utilized directly in EONs.

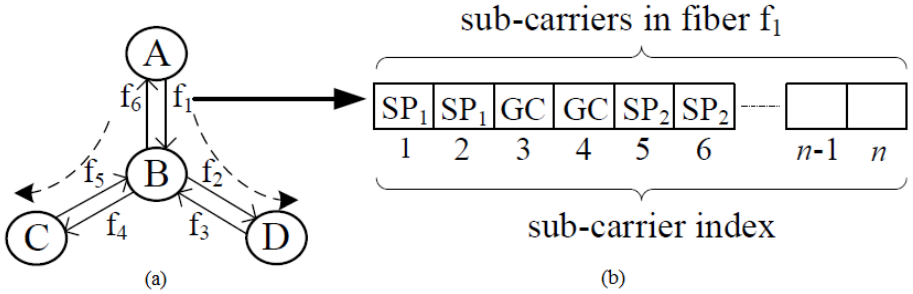


Figure (3.1) An example of RSA in EONs.

The routing and spectrum allocation (RSA) algorithm in EONs allocates a spectrum path (SP) between two optical nodes. The SP should be assigned the same spectrum along its routing path while the sub-carriers allocated along that path should be contiguous. The former is known as the *spectrum-continuity* constraint while the latter is known as the *sub-carrier consecutiveness* constraint. Further, when two SPs share one or more common fiber, the corresponding allocated sub-carriers have to be separated by the guard-carriers (GC). The size of GC is not trivial and may be in the order of one or multiple sub-carrier(s) [19]. An example of RSA is shown in Figure 3.1 where Figure 3.1(a) is a star network with 2 directional fibers per link and GC = 2. In the network shown in Figure 3.1(a), there are two SPs, SP_1 of 2 sub-carriers from node A to node C, and SP_2 of 2 sub-carriers from node A to node D. Figure 3.1(b) shows the spectrum allocation in fiber f_1. Each sub-carrier in

fiber f_1 has an index and there are n sub-carriers in total. Sub-carriers 1 and 2 are assigned to SP_1 whereas sub-carriers 5 and 6 are assigned to SP_2. Since the sub-carriers within SP_1 are consecutive, no GC is required. However, to be able to separate SP_1 from SP_2 in fiber f_1, GCs are necessary. Hence, sub-carriers 3 and 4 cannot be assigned to SP_2. Since the size of GC is 2, fiber f_1 requires 6 sub-carriers to accommodate SP_1 and SP_2 as shown in Figure 3.1(b).

The RSA problem can be formulated in a number of ways with various objectives under various assumptions. Broadly speaking, the RSA problem in EONs can be classified into two broad versions : offline RSA, whereby the topology of the network and the traffic demands are known in advance, and online RSA, where the network topology is known in advance but the traffic demands arrive in some random order. Since RSA is a key to efficient spectrum allocation in EONs, the offline RSA and the online RSA are explained in detail in the following section.

3.2 Offline RSA

The input to the offline RSA problem consists of a set of traffic demands and the network topology, both of which are known in advance. The output is an SP and sub-carrier allocation along this path with the objective to minimize the total amount of utilized spectrum or the maximum sub-carrier index (MSI) (either over the whole network or on any link). Offline RSA is a well-researched problem in the literature. The authors in [20] computed the complexity of the offline spectrum allocation problem. Using results from the graph coloring theory, it was shown that the spectrum allocation problem is NP-hard. Since the RSA problem is NP-hard, a variety of integer linear program (ILP) formulations have been proposed for the offline RSA in EONs, each dedicated to solve a specific problem variant. However, these ILP formulations cannot be solved within a reasonable amount of time for problem instances involving larger network topologies. Hence, a number of heuristic algorithms have also been proposed to solve the offline RSA problem in larger networks.

The authors in [6], [21], [22] proposed a Link-based ILP formulations of offline RSA as

a multi-commodity flow problem. For the RSA, the authors also considered *guard-carrier* constraints while allocating sub-carriers to the traffic demands in addition to the *spectrum-continuity* and *sub-carrier consecutiveness* constraints with the objective to minimize three factors: the maximum number of sub-carriers allocated on any fiber, the maximum sub-carrier index allocated on any fiber, and the total number of sub-carriers over all fibers.

The authors also calculated the upper and lower bounds on the optimal solution for the ring and mesh networks under both predetermined and non-predetermined routing. The lower bounds were obtained using cut-set techniques and the latter was shown to provide tighter bounds. It was also shown that, in ring networks with uniform demands, the lower and upper bounds are tight. Finally, two heuristic algorithms were proposed. The first algorithm, referred to as shortest path with maximum spectrum reuse (SPSR), uses shortest path routing and the first-fit spectrum allocation strategy to assign sub-carriers to demands in the decreasing order of their size. The second algorithm referred to as balanced load spectrum allocation (BLSA) considers the k shortest paths as candidates for each connection demand, and then selects the one which minimizes the maximum link load to balance the use of spectrum across the network links.

The authors in [23], [24] proposed a similar link-based formulation of the RSA problem known as the routing, wavelength assignment, and spectrum allocation (RWSA) problem with the objective to select a set of line rates for each demand and a corresponding set of lightpaths to minimize the use of total spectrum in the network. Another version of path-based ILP formulation of the RSA problem was presented in [25]. The authors used a path-based formulation with the objective to minimize the number of sub-carriers that are used in any links of the network. The authors used the ILP formulation to assign one of the predetermined paths to each demand, while also satisfying the *spectrum-continuity* and *non-overlapping spectrum* constraints. This path-based formulation is more compact than the link-based ILP of [6]. However, the number of decision variables and constraints are huge that it cannot be solved directly. Accordingly, a heuristic algorithm called adaptive frequency assignment with collision avoidance (AFA-CA) was proposed to select the path

for each demand. The authors in [7], [26] presented another path-based ILP formulation for the RSA problem where the objective was to minimize the maximum sub-carrier index assigned on any link in the network. The concept of a channel as a set of contiguous sub-carriers of a given width to model the *spectrum-continuity* constraint was introduced in [27]. For a given spectrum width t, all possible channels with t sub-carriers are defined on each link of the network. Then, the RSA problem is transformed to one of routing and channel allocation, in which channel assignment implies allocation of contiguous spectrum, and no explicit *spectrum-continuity* constraints are required. This results in a more compact formulation that achieves a significant speed-up in running time compared to those that directly account for the *spectrum-continuity* constraints [27].

3.3 The online RSA

The online RSA in EONs is even more challenging due to the random traffic arrival/departure and the fluctuation of the traffic demands over time. Depending on the network state, the available spectral resources may or may not be sufficient to establish a connection. The network state consists of the physical links and the spectrum assignment for all active connections. As the network evolves, a current optimal routing algorithm might no longer provide the optimal spectral utilization over time. Thus, each time a new connection request arrives, an algorithm must be executed in real time to determine whether it is feasible to accommodate the new connection request, and, if so, to perform the RSA. If a new connection request cannot be accommodated, then it is blocked. Thus, the blocking probability of connection requests arises as the key objective in an online RSA algorithms.

Since the online RSA problem is more challenging, most of the studies propose heuristic algorithms to solve the online RSA problem. Based on whether the online RSA algorithm can solve the RSA problem jointly (i.e., in one step) or separately, (i.e., in two steps), it can be broadly classified into one-step algorithms and two-step algorithms.

3.3.1 One-step algorithms

One-step online RSA algorithms solve the RSA problem simultaneously but sub-optimally, typically using a greedy approach. In [28], the author considered a dynamic version of the RWSA problem and decomposed it into a rate selection problem and a dynamic routing and channel selection problem. The dynamic routing and channel selection problem was solved by representing the state of the network using an auxiliary graph. Two different one-step online RSA algorithms were introduced in [29]. In the first, a modified Dijkstra shortest path algorithm was employed that used the links with enough sub-carriers in the increasing order of the path length. The second one-step RSA algorithm builds a path vector tree routed at the source node such that all links along each path have sufficient contiguous spectrum for the request. The algorithm then searches the path vectors to find the link with sufficient available spectrum and minimum cost. A modified Dijkstra algorithm similar to the one in [29] was proposed in [30] to solve the online RSA problem in one step. The main difference is that the spectrum availability along the path is maintained by tracking of the channels as defined in [27] that may be sufficient to accommodate the given connection request.

3.3.2 Two-step algorithms

Two-step online RSA algorithms decompose the problem of RSA into two sub-problems, the routing problem and the spectrum assignment problem, which are then, solved sequentially.

The routing algorithm may be static, where the paths are computed and fixed, or adaptive, where the computed paths and their order may vary depending on the state of the network. In [29], the authors proposed routing algorithms that computed multiple alternate paths for each connection requests. Then, the spectrum allocation algorithm determines the set of available contiguous slots are assigned to a request. A first-fit policy [29], [31] selects the lowest indexed sub-carrier, a random-fit policy [32] randomly allocates one of the available sub-carriers, whereas a best-fit policy selects the smallest set of sub-carriers that

can satisfy the connection request. The authors in [33] proposed an improvement in the operation of the first-fit policy. This study proposed an evolutionary algorithm to search for the most feasible spectrum ordering for first-fit so as to minimize the blocking probability.

The study in [31] investigated the optimal slot width for EONs by measuring the blocking probability under dynamic traffic using Monte Carlo simulations. Each demand was routed on its shortest path, and the first-fit policy was used for spectrum allocation. The authors in [34] proposed an adaptive routing algorithm to solve the first sub-problem. Specifically, routing tables at each node associate a probability with each pair. A dynamic ant colony optimization (ACO) algorithm is executed continuously that updates these probabilities based on spectrum utilization. Once the path has been determined, the first-fit policy is used for spectrum allocation.

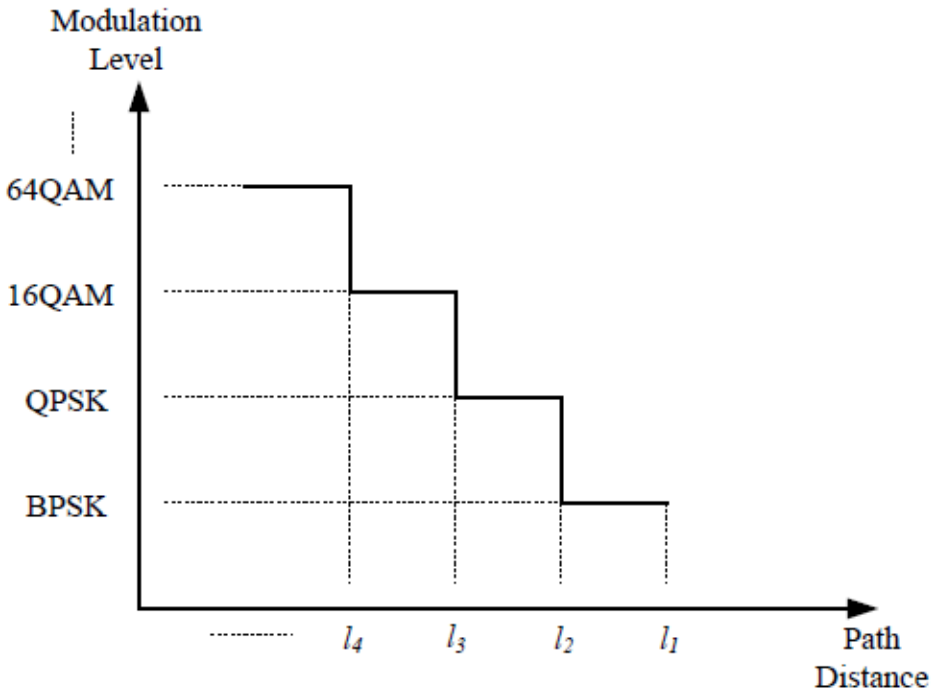


Figure (3.2) Modulation level as a function of the transmission distance

3.4 Modulation-aware Distance adaptive RSA

Since EONs uses OFDM as the modulation format, EONs has further flexibility of choosing the number of modulated bits per symbol for each sub-carrier based on the transmission distance. In particular, each OFDM sub-carrier can be modulated individually using, for example, single bit per symbol binary phase-shift keying (BPSK), 2 bits per symbol quadrature phase shift keying (QPSK), 3 bits per symbol quadrature amplitude modulation (QAM) etc [7] to account for link impairments. This allows the demands with the same data rate to be allocated different amounts of spectral resources depending on the quality of their path [35], [36], [37], [38], [39], [40]. In contrast, in traditional WDM networks, changing the modulation level/format requires the use of a different transponder [41]. Thus, given the length of an optical path, higher modulation level can be used over the path as shown in Figure 3.2. Therefore, transmissions over shortest optical paths are able to utilize higher modulation levels, which can carry more bits per symbol and results in lower number of sub-carriers.

Distance-adaptive spectrum allocation was first introduced in [42] where the authors evaluated the trade-off between the width of the spectrum and the transmission reach to improve spectrum utilization [43] by adaptively choosing the modulation format based on the level of impairments. A high-level modulation format with narrow spectrum and low signal to noise ratio (SNR) tolerance may be selected for a short path, whereas a low level modulation with a wider spectrum and high SNR tolerance may be used for a longer path [44]. The authors in [7], [45] studied the offline version of the distance-adaptive RSA problem of finding the appropriate modulation level based on the transmission distance to create a unique Routing, Modulation Level and Spectrum Allocation (RMLSA) problem. The authors also presented a decomposition method that broke RMLSA into its two substituent sub-problems, namely, (i) routing and modulation level (RML), and (ii) spectrum allocation (SA), and solves them sequentially. Most of the research work in distance-adaptive RSA considers the online version of the problem where the heuristic algorithms attempt to accommodate

randomly arriving connection requests. Several algorithms follow a two-step approach similar to the online RSA. The authors in [39], [42] presented the two-step algorithms for the online distance-adaptive RSA problem. These algorithms compute a number of fixed-alternate routing paths (using static routing) for each source-destination pairs, and order them in terms of decreasing path length. In the second step, these algorithms employ the first-fit spectrum allocation and consider each path sequentially to find a number of contiguous sub-carriers can accommodate the requested data rate over the given path length. A similar algorithm is used in [38] for routing of super-wavelength demands.

The authors in [44] studied the online distance-adaptive RSA problem referred to as dynamic impairment-aware RSA (IARSA). The authors studied two variants of IARSA, one in which the modulation format of the incoming signal may be modified by the regenerators, and the other in which the modulation format remains constant. The authors in [46] proposed the quality of transmission (QoT) aware online RSA algorithm that consists of three stages: path calculation, path selection, and spectrum assignment. The Dijkstra algorithms and k-shortest path algorithms were utilized for computing paths, while fiber impairments and non-linearity effects at the physical layer were modeled using a closed-form expression that is used to estimate the QoT along a given path. Then, for each request, the most feasible path is chosen with respect to the optical signal-to-noise ratio (ONSR), as determined by the physical layer model

3.5 Chapter Summary

In this chapter, we present the routing and spectrum allocation (RSA) algorithm in EONs. The RSA in EONs has to follow *sub-carrier consecutiveness*, *spectrum-continuity* and *guard-carrier* constraint while allocating a spectrum path to a connection demand. The RSA in EONs is NP-hard. Hence, various Integer Linear Programming (ILP) formulation and heuristics algorithms are proposed for RSA in EONs. Based on the network topology and the traffic matrix, these algorithms can be further classified into offline RSA, whereby the topology of the network and the traffic demands are know in advance, and online RSA

where the network opology is known in advance but the traffic demands arrive in random order. Since, EONs uses OFDM as the modulation format, EONs has the further flexibility of choosing the number of modulated bits per symbol for each sub-carrier based on the transmission distance. This gives rise to distance-adaptive RSA algorithms in EONs.

PART 4

SPECTRAL FRAGMENTATION IN ELASTIC OPTICAL NETWORKS

4.1 Introduction

In EONs, the RSA algorithm has to consider the *spectrum-continuity* constraint while allocating a SP to a connection demand. As the connection demands arrive/depart randomly, the spectral resources can get fragmented as the same block of spectrum may not be available along all the links of a path despite each link having sufficient bandwidth. Fragmented spectral resource may cause significant amount of connection blocking in EONs. While, the *spectrum-continuity* constraint is the major cause of spectrum fragmentation in EONs, the spectral resources may also get fragmented due to the variable data rates [47]. Since, fragmentation due to the *spectrum-continuity* constraint is more profound in EONs, in this chapter, we explain the spectrum fragmentation in EONs due to the *spectrum-continuity* constraint and also highlight some of the related works in this issue. Further, in this chapter, we also propose a novel defragmentation algorithm to mitigate spectrum fragmentation and consolidate fragmented spectral resources [9] and a novel multipath RSA algorithm [48] that modifies the existing RSA algorithm to cope with the fragmented spectral resources. Simulation results and the summary are presented at the end of the chapter.

4.2 Fragmentation Problem in EONs

In EONs, the RSA algorithm has to consider the *spectrum-continuity* and *sub-carrier consecutiveness* constraints while assigning a spectrum path (SP) to any incoming connection demands [6]. Due to these additional constraints in EONs, the spectral resources in EONs can get fragmented due to the random traffic arrival/departure and the fluctuation of the traffic demands over time [8]. In EONs with dynamic traffic, the frequent set-up and tear down of connections can lead to significant fragmentation of spectral resources [8]. As the spectral

resources are allocated and deallocated dynamically and non-uniformly, spectrum vacuums are created by the outgoing *SPs*. The spectrum vacuums, if not properly managed, will accumulate over time and become unstable. Due to the *spectrum-continuity* and *sub-carrier consecutiveness* constraints in EONs, several spectrum slots in between connections remain unused as the white spaces shown in Figure 4.1. These spectral fragments are small, scattered and may not be enough to establish new connections because of aforementioned constraints in EONs. As a result, the spectral fragments may increase the maximum used sub-carrier index (MSI) in each fiber or decrease the probability of finding sufficient contiguous sub-carriers for new connections. In a broad sense, fragmentation include any sub-optimal spectrum distribution that causes unnecessary rejection of the incoming connection demands even when enough spectral resources exist scattered in the network.

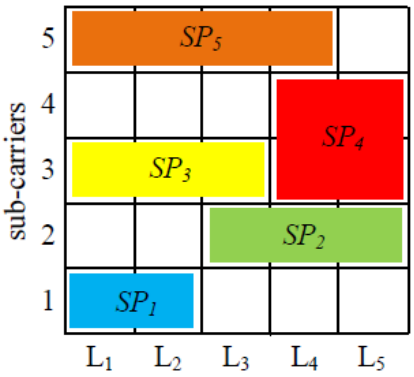


Figure (4.1) (a) Spectral defragmentation in networks with 5 *SPs* along 5 consecutive links

With spectral fragmentation in EONs, new incoming connection requests are then forced either to utilize more spectrum in the fiber or are blocked even though sufficient spectrum are available. Hence, the spectral fragments represent the wastage of the expensive spectral resources which may lower spectral usage and increase blocking. The example of spectrum fragmentation shown in Figure 4.1(a) shows the conditions of spectrum slots in 5 consecutive links - L_1, L_2, L_3, L_4 and L_5. Each link has 5 sub-carriers and 5 *SPs* are established along these links. The white spaces in between these *SPs* represent unused spectrum fragments.

Any new connection that requires more than one sub-carrier cannot be established along any of these links.

4.3 Spectral defragmentation schemes

Spectrum defragmentation is the process of consolidating the spectrum in order to free up the spectrum for future connection requests which otherwise may not be usable. The spectral defragmentation in EONS was first introduced in [8] and was demonstrated experimentally in [49]. Most defragmentation schemes can be categorized as either *proactive* or *reactive*. Proactive schemes are carried out independent of whether there is any blocking in the network due to spectral fragmentation, and are executed periodically most of the time. The goal of the proactive schemes is to reduce the spectrum fragmentation measured by some network-wide indicators such as fragmentation index [47], fragmentation ratio [50], [51], utilization entropy [52] etc. Reactive schemes, on the other hand, are triggered when the connection requests are rejected due to the insufficient spectral resources. The primary goal of the reactive schemes is to rearrange existing spectrum fragments to make room for the blocked connection request. If enough sub-carriers are not found, then the connection request is blocked.

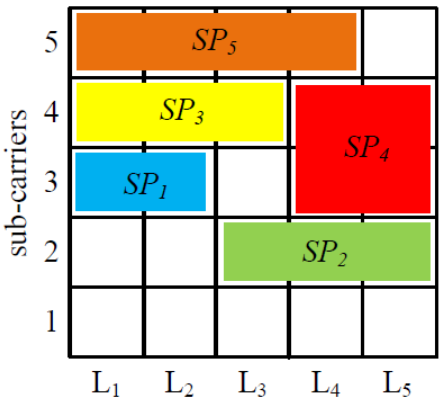


Figure (4.2) An example of the Push-pull defragmentation

4.3.1 Proactive defragmentation schemes

The authors in [8] explored proactive defragmentation schemes in EONs with the goal to minimize the number of network reconfiguration. The heuristic algorithms proposed in this paper utilized both re-routing and spectrum reallocation in a make-before-break fashion. The authors in [53] proposed a proactive scheme that reserves a block of contiguous sub-carriers for each source-destination pair. In addition, the sub-carriers that are not reserved may be shared on demand among all connections. In this study, the author assumes that the demands vary with time, but as long as a demand stays within its reservation, it can always be accommodated. The authors in [51] introduced two proactive algorithms that allocates spectrum based on the current state of the fragmentation in the network. While one algorithm assigns each new connection request to a path that minimizes a network-wide fragmentation ratio defined in the same study, the second algorithm attempts to utilize sub-carriers that are already used the most in the network.

4.3.2 Reactive defragmentation schemes

Reactive defragmentation schemes employ defragmentation techniques to restore the network's ability to accommodate connections that would otherwise be blocked. The objective of defragmentation is to rearrange the scattered spectrum of existing traffic demands so as to consolidate available spectrum into large continuous blocks that may be used to establish future requests. Reactive defragmentation schemes may be broadly classified as periodic or path-triggered [54]. Periodic defragmentation schemes are initiated periodically at regular intervals or whenever the fragmentation metrics exceeds certain thresholds while path-triggered defragmentation schemes are initiated when the new connection requests get blocked. In [39], the authors proposed reactive defragmentation schemes where existing optical paths that conflict with the candidate route of the blocked request are subjected to rerouting and spectrum reallocation. One of the hitless defragmentation technique known as Push-pull technique was proposed in [55]. Push-pull defragmentation technique allows established SPs to shift along free and contiguous spectrum bands in a hitless fashion, if the

established SPs does not changes its route and the SPs does not transgress other established SPs. In Figure 4.1, SP_3 can be shifted from sub-carrier index 3 to 4 and SP_1 can be shifted from sub-carrier index 1 to 3 to consolidate the spectrum at the upper end of the spectrum as shown in Figure 4.2 in a hitless manner. Here, the shifting of SPs take place one step at a time. Hence, consolidating the SPs at the upper end of the spectrum requires 3 steps. However, SP_1 cannot be directly shifted from sub-carrier index 1 to 4 as it cannot transgress the SP_3 at sub-carrier index 3. In case more than one SPs are shifted, parallelism, can be exploited to reduce the overall reconfiguration time. A path-triggered defragmentation strategy was proposed and investigated in [54]. When a new request arrives, an online RSA algorithm is first executed to accommodate it. If a path with sufficient spectrum cannot be found, then a defragmentation algorithm is triggered in the candidate paths to accommodate the new connection request. The spectral defragmentation algorithm proposed in [56] tries to minimize the number of connections that would be affected to accommodate the blocked request.

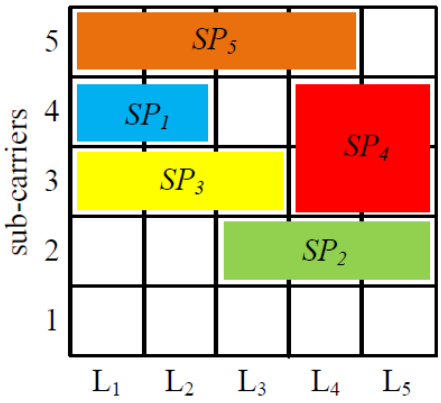


Figure (4.3) An example of the Hitless defragmentation

The authors in [57] proposed a defragmentation scheme exploiting a new technique for fast wavelength tracking in coherent receivers. In the proposed defragmentation technique, already established SPs are capable of hopping over already existing SPs. This technique can be applied to a single-carrier connection or each of the sub-carriers forming a super-

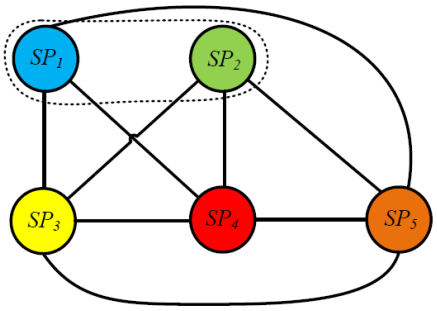


Figure (4.4) An auxiliary graph

channel. An example of this defragmentation scheme is shown in Figure 4.3 where SP_1 is moved to sub-carrier index 4 by moving over SP_3 directly. This technique is faster and more efficient than the technique in [55] since it does not rely on sequential tuning of wavelengths. However, all these defragmentation schemes do not maximize the consolidation of scattered spectral resources. In the following section, we present a proactive defragmentation technique which does the complete defragmentation of spectral resources using Independent sets [9].

4.3.3 Spectral Defragmentation Using Independent Sets

One obvious way to reduce the spectral fragmentation is to reconfigure existing connections with the goal of consolidating the spectrum allocation. The scattered and fragmented spectrum slots can be consolidated by either shifting the existing sub-carrier allocation between one node pair to a different group of sub-carriers, assigning a new route to an existing connection, or by both while maintaining the *sub-carrier consecutiveness* and *spectrum-consecutiveness* constraints. Network administrators can perform the spectrum defragmentation on a periodic manner to consolidate spectrum or on demand when the links' sub-carrier index increases beyond a threshold (indicating the potential blocking of future connection requests).

In this section, we propose an efficient heuristic to consolidate the spectrum allocation using an auxiliary graph with the primary objective of reducing the MSI in each link and thus improving spectrum utilization and reducing connection blocking. Assume $SP_i = < p_i, t_i >$ to be a connection where p_i is the path and t_i is the request size (in terms of the number of

sub-carriers). We construct an auxiliary graph based on the existing spectrum allocations using the following steps :

1. Combine two or more spectrum paths with same source and destination into one spectrum path.

   $SP_{ij} < p_{ij}, t_{ij} >= SP_i < p_i, t_i > \cup SP_j < p_j, t_j >$ where $p_i = p_j, p_{ij} = p_i$ and $t_{ij} = t_i + t_j$

2. Create a vertex for each spectrum path.

3. Create an edge between two vertices i and j if $t_i \neq t_j$ or vertices i and j cannot be assigned sub-carriers with the same index.

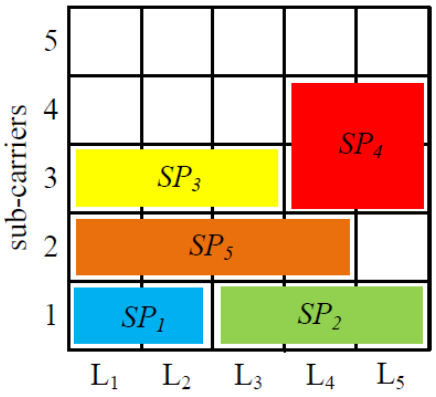


Figure (4.5) Sub-carrier allocation after defragmentation

For example, as shown in Figure 4.4, there is no edge between vertices SP_1 and SP_2 as both of these *SPs* require 1 sub-carrier and they can both be assigned using sub-carriers with the same index. After constructing the auxiliary graph, we can use Algorithm 1 which uses independent sets to consolidate the spectrum allocation. Algorithm 1 uses a greedy approach to partition the given *SPs* into different disjoint sets. First, the maximum independent sets (MIS) is found and the *SPs* which belong to the MIS are grouped together into one set and these vertices are deleted from the auxiliary graph. Then another MIS is found and the *SPs*

are grouped accordingly. This process continues until all the *SPs* belong to an unique set. Using Algorithm 1 in Figure 4.4, we can divide the *SPs* into 4 disjoint sets - (SP_1, SP_2), (SP_3), (SP_4), (SP_5). Figure 4.5 shows the resulting sub-carrier assignments after applying Algorithm 1 on the auxiliary graph in Figure 4.4. Among these sets - (SP_3), (SP_4), (SP_5), SP_3 is assigned first as the hop length is 4 which is highest among these sets.

Algorithm 1 Spectrum Defragmentation using Independent Sets

1: Transform the *SPs* to an auxiliary graph.
2: Group the vertices into different disjoint sets according to the MIS found. Find MIS and the *SPs* that belong to the MIS are grouped. Remove these *SPs* from the graph. Find another maximum IS and group *SPs*. Continue this, until all the *SPs* belongs to an unique set.
3: Order the sets in the decreasing order of the hop length.
4: Assign the first available sub-carriers to the connections with the longest hop length from Step 3.
5: Repeat Step 4 until all the connections are established.

Intuitively, an *SP* with the longer hop length should be assigned with sub-carriers having lower index. This will help consolidate the spectrum allocation, lower the MSI in each fiber link and leave less empty unused spectrum slots in between connections. Having connections with least hop length at the top of the spectrum also makes it easier to route those connections through alternate paths in order to make way for bigger connections if necessary as connections with less hop length may impact less on spectrum assignments along the alternate route.

4.3.4 Performance Evaluation of the Proposed Defragmentation Algorithm

The authors in [39] proposed a dynamic algorithm that assigns routes and sub-carriers between two optical nodes in networks with dynamic traffic. We use this algorithm to assign routes and sub-carriers in networks with dynamic traffic and use the proposed spectrum defragmentation algorithm periodically to consolidate the scattered spectrum. We then compare the performance of the case with periodic defragmentation to the case without periodic defragmentation in terms of the blocking probability and the MSI used in any

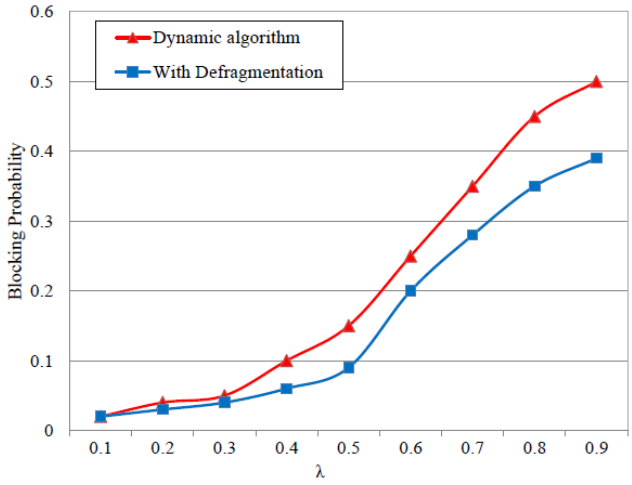


Figure (4.6) λ vs. blocking probability

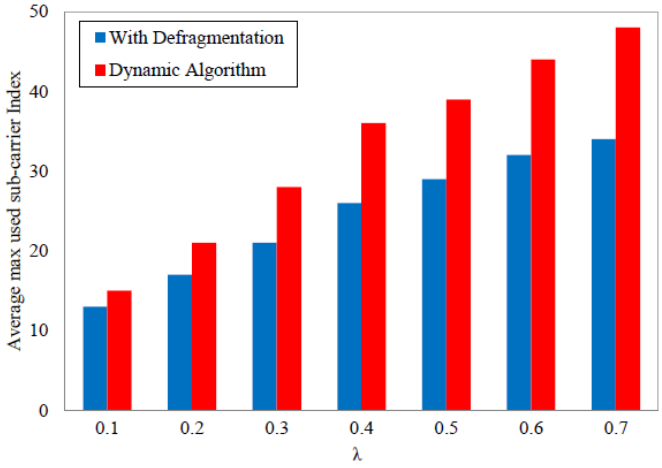


Figure (4.7) λ vs. maximum used sub-carrier index in any fiber

fiber. Connection requests arrive according to the Poisson process at a rate of λ in a 14-Node NSF network. The holding time of each connection request conforms to a negative exponential distribution with parameter μ. There are 50 sub-carriers in each fiber link. Blocking probability is calculated as the ratio of number of the connection requests that are blocked to the total number of connection requests in the network.

Figure 4.6, it is clear that proposed heuristic can significantly lower the blocking probability. The reduction in blocking probability is more significant with heavier traffic load (i.e., larger λ). The blocking probability is almost the same for $\lambda = 0.3$ as there are less number connections and huge number of sub-carriers. Hence, even though there are spectral fragments, more sub-carriers are available to establish further connection requests. However, from onwards, the proposed heuristic can result in less blocking probability as the proposed heuristic consolidates the spectrum usage and provide new connections more spectrum to use. We also compare the MSI used in any fiber when increasing λ as shown in Figure 4.7. The proposed heuristic yields lower the MSI in any fiber. The difference is even higher for higher λ as the spectral fragmentation will be ever worse for heavier traffic. When $\lambda = 0.7$, the proposed heuristic can lower the MSI by about 30 %.

4.4 Multipath routing to cope with spectral fragmentation

While a single routing path is trivial to compute and easy to setup, a single routing path may not be the most efficient as it may not be able to find the contiguous sub-carriers along the routing path because of the spectrum fragmentation leading more connection requests to be blocked. Multipath routing provides more options to be explored by utilizing more than one SP to satisfy a connection demand. However, splitting a single path over multiple paths could increase the required sub-carriers as more guard-carriers are needed in each path. Figure. 4.8 shows a simple network with five nodes where there are five sub-carriers in each links. Consider the size of GC is 1 sub-carrier. If there is a request R_1 of two sub-carriers between nodes A and C, then this request has to be blocked since it requires 3 sub-carriers (including 1 GC) to satisfy the demand but there are only 2 sub-carriers available. However, with multipath routing, this connection request can be satisfied by splitting the demand of two sub-carriers into two demands of one sub-carrier each. This reduces the blocking probability; however note that the 2 GC are required in this case. EONs support various modulation formats. The number of bits per symbol and the filter bandwidth can be adjusted adaptively based on the distance to provide a distance adaptive spectrum

allocation scheme [42]. Up to now, spectrum allocation in optical OFDM networks has used single path routing with a uniform modulation format in the entire network. However, the uniform modulation assignment cannot fully utilize the available spectrum, due to the fact that optical transmission systems are commonly designed with consideration of the worst case scenario in terms of transmission performance [42].

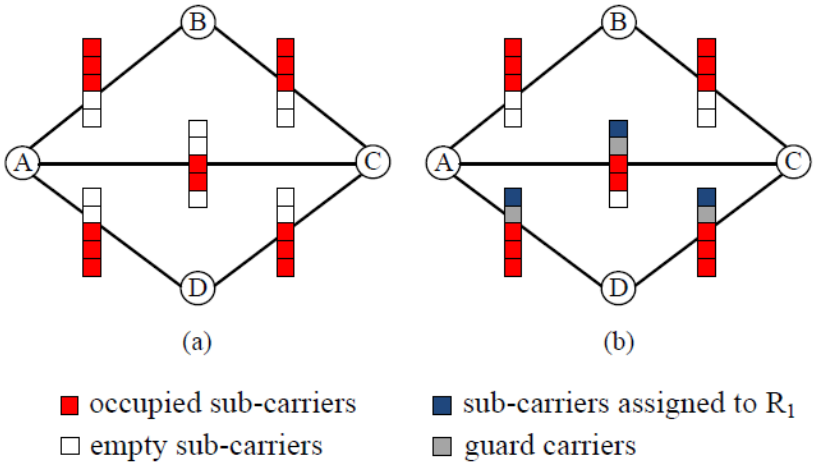


Figure (4.8) (a) An example Network (b) RSA using multipath routing

4.4.1 Related Works in Multipath Routing

RSA in EONs has been investigated under both static and dynamic traffic conditions as we have seen in previous chapter. In the most existing works, the path length and the modulation format are not taken into consideration and the uniform modulation format is utilized for the entire network. Hence, the modulation format suitable to transfer the data along the longest path is selected and is applied to the entire network. However, many paths in the network are much shorter than the longest path and can support higher modulation formats. Using a single modulation format can result in inefficient use of the spectrum. In EONs, the filter bandwidth and the number of bits per symbol can be adjusted adaptively based on the transmission distance [42]. Distance-adaptive spectrum allocation can improve the spectrum efficiency by adaptively selecting the suitable modulation formats for different paths depending on the path distance. The low data rate modulation format can

be selected for longer paths and high data rate modulation format can be selected for shorter paths. RSA with distance-adaptive approach was first proposed in [42] where the authors studied the optical filter width and modulation format to determine the available spectrum resources required. The authors also proposed the algorithm where k-shortest paths are computed and the required spectrum is determined depending on the path distance and the lowest spectrum slot index is selected. Recently, multiple paths have been widely suggested for routing requests in EONs [58], [50], [59], [60], [61], [61]. Two heuristic algorithms for multipath routing were proposed in [58]. In the first algorithm, path selection and spectrum assignment are done together whereas in the second algorithm, the paths are pre-computed with respect to their differential delay. In both algorithms, the available spectrum (total of all fragments) is aggregated starting from the shortest path until the request is fulfilled. However, in these algorithms, as the SPs are reserved in a succeeding manner, the last SP may not be fully utilized. Moreover, the number of *SPs* and guard-carriers required may increase. An algorithm that uses a hybrid single-/multi-path routing (HSMR) scheme was presented in [50]. The algorithm considers a set of candidate paths in the order of increasing weight, and takes the fragmentation into account as well. The authors in [59] presented the dynamic multi-path routing scheme that selects the candidate paths such that the maximum differential delay along these paths does not exceed a given threshold. An offline version of the multi-path RSA problem was studied in [62]. The authors formulated the multipath routing as a path-based ILP problem. In addition to the usual *spectrum-continuity* and *sub-carrier consecutiveness* constraints, this ILP also imposes differential delay constraints on the multiple paths for a given connection. An offline multi-path provisioning scheme for OFDM based EONs was first introduced in [61]. The problem, referred to as the "static survivable multi-path routing and spectrum allocation" (SM-RSA), was first formulated as a path-based ILP. An online version of the same problem is studied in [63] and again an ILP formulation and a heuristic algorithm are presented.

In this section, we propose a novel multipath RSA algorithm that uses minimal number of routing paths with distance-adaptive modulation called Modulation-aware Multipath

Routing Algorithm and Spectrum Allocation (MMRSA) algorithm. MMRSA allocates the available spectral resources efficiently by minimizing the number of routing paths required, which in turn reduces the number of guard-carriers. With minimal number of routing paths or *SPs*, we may also mitigate the spectrum fragmentation issue, while making full use of the OSNR margin on each spectrum path by adaptively assigning modulation format based on the transmission path length.

4.4.2 Modulation-aware Multipath Routing Algorithm

Preliminaries An EONs can be represented as an undirected graph, $G(V, E, N)$, where V represents the set of nodes, E represents the set of fiber links and N represents the number of sub-carriers in each fiber. The length of a fiber link, $e \in E$ is denoted as LD_e. The length of an SP is the sum of the lengths of all fiber links that it traverses, which is defined as $L_{SP} = \sum_{e \in SP} LD_e$. The set of modulation formats that can be supported by the network are denoted as $M_F = \{m_{f_1}, m_{f_2}, ..., m_{f_n}\}$ where $m_{f_1} \geq m_{f_2} .... \geq m_{f_n}$. For each modulation format, there exists a maximum transmission distance, denoted as L_{m_k} [1]. A modulation format m_{f_k} can be applied to a spectrum path SP, only if $L_{SP} \leq L_{m_k}$. A connection request is represented as $R(s, d, n)$, where s, d and n represents source, destination and the number of sub-carriers of highest modulation format requested, respectively. However, if the transmission distance between s and d is greater than transmission reach of the highest modulation format, then the modulation factor M_f has to be applied to determine the effective number of sub-carriers required as follows:

$$M_f = \left\lceil \frac{b_h}{b_k} \right\rceil \tag{4.1}$$

where b_h is the bit rate per symbol of a highest modulation format and b_k is the bit rate per symbol of the selected modulation format. Then the total number of sub-carriers ($\bar{n}$) required for R can be calculated as follows:

$$\bar{n} = M_f \times n \tag{4.2}$$

Modulation-aware Multipath Routing and Spectrum Allocation (MMRSA) Algorithm The basic idea of Modulation-aware Multipath Routing and Spectrum Allocation (MMRSA) algorithm is to first identify multiple candidate SPs sequentially starting with the shortest path. Once the required number of candidate SPs is found, MMRSA does not allocate the sub-carriers in these SPs right away. Instead, MMRSA tries to identify the minimum number of SPs needed for the connection demand by trying the various combinations of the SPs. The details of MMRSA are given in Algorithm 2.

In MMRSA, k shortest paths from s to d are computed in line 1. In line 2, the modulation format for each of the k paths are determined based on their transmission distance and the corresponding modulation factor (M_f) is calculated using Eq. (4.1) and (4.2). In line 3, the maximum number of contiguous available sub-carriers available in each path is computed. Using this, the available number of sub-carriers is computed by subtracting the guard-carrier in line 4. In line 5-15, the number of paths required to satisfy the connection request R are computed in ascending order of the path distance and stored in the set UB. If the total capacity of these paths (usable sub-carriers/M_f) is less than n, then the request R is blocked in line 17. Otherwise, the best combination of paths that will serve the request R is computed in line 19-29. For this, the last path $p_l \in UB$ is removed and stored in P_0 in line 20. Then, all paths in UB are arranged in the descending order of their capacity in line 22. Since, p_l will be there in the final solution, no matter which combination is chosen, different combination of SPs are checked with p_l to find the minimum number of SPs. For example, only p_l is considered first, and then $p_l + p_1$ is considered and so on. This process is continued until the $|UB|$ combinations are reached (i.e. $p_l + p_1 + p_2 + \cdots + p_{|UB|}$). Note that each combination is stored in set P_i as shown in line 25. In line 31, the combination that requires the lowest number of sub-carriers is selected. Lower number of paths is used to break the tie if any combinations require the same number of sub-carriers. Note that the last combination ($p_1 + p_2 + \cdots + p_{l-1} + p_l$) is the same as that used in the multipath algorithm proposed in [58]. In the proposed MMRSA algorithm, more combinations are explored to reduce the number of sub-carriers (as well as number of paths). Hence, the performance

of the proposed algorithm is expected to be equal or better than the multipath algorithm evaluated in [58]. By arranging all the path in UB in descending order of their capacity, MMRSA only needs to check $\mathcal{O}(n)$ path combinations instead of $\mathcal{O}(n^2)$ combinations. Hence, the MMRSA algorithm is computationally less expensive and utilizes less number of SPs than the multipath algorithm proposed in [58].

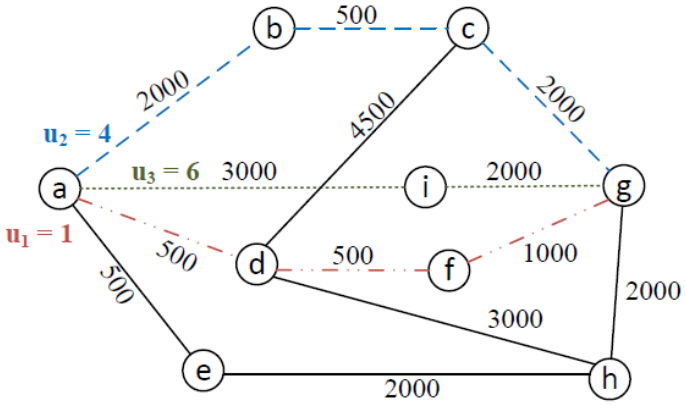


Figure (4.9) An 9-node network

An Example of the Algorithm Fig. 4.9 shows an example of the network where the demand (D_m) requires 4 sub-carriers of the highest modulation format from source a to destination g. Here, u_i is the usable number of sub-carriers (i.e. maximum contiguous sub-carriers minus guard-carrier) in path p_i. With $k = 3$, there are three shortest paths from a to g: $a - d - f - g$ (path 1, length = 2,000), $a - b - c - g$ (path 2, length=4,500), and $a - i - g$ (path 3, length=5,000). Suppose that the transmission range for the highest and the second highest modulation format is less than 3,000 and 6,000 units resp. and the spectrum required for the second highest modulation format is twice as compared to the highest one, i.e. ($M_f = 2$). Therefore, the modulation format ratio for paths 1, 2, and 3 are $M_1 = 1$, $M_2 = 2$ and $M_3 = 2$, resp. Here we consider the demand in number of sub-carriers corresponding to the highest modulation rate format. In this example, the size of the guard-carrier is 1 sub-carrier.

Path 1 has only one usable sub-carrier ($u_1 = 1$) while Path 2 has 4 ($u_2 = 4$). As the

Algorithm 2 Modulation-aware Multipath Routing and Spectrum (MMRSA) Allocation Algorithm

Input: Network $G(V, E, N)$, a connection requests $R(s, d, n)$
Output: Route and spectrum allocation for $R(s, d, n)$

Compute k shortest paths from s to d
Select modulation format for each path based on its distance.
For each path, calculate the maximum consecutive sub-carriers available using Eq. (4.1) and (4.2).
Calculate the usable sub-carriers $(u_1, u_2, , u_k)$ in k paths by subtracting guard-carrier from the available sub-carriers.
$UB \leftarrow 0$
$S_c \leftarrow 0$
for each path i in k **do**
    $S_c \leftarrow S_c + u_i/M_f$
    if $S_c < n$ **then**
        $S_c = S_c + u_i$
        $UB \leftarrow p_i$
    else
        break;
    end if
end for
if $|S_c| < |n|$ **then**
    Block R
else
    $i \leftarrow 0$
    $P_i \leftarrow p_l$
    $UB \leftarrow UB \setminus p_l$
    Sort the set UB in descending order of their demand.
    while $i \leq |UB|$ **do**
        $i \leftarrow i + 1$
        $P_i \leftarrow p_i \cup P_{i-1}$
        if P_i satisfies R **then**
            Record the number of sub-carriers required in P_i
        end if
    end while
end if
Allocate proper sub-carriers in P_i to satisfy R.

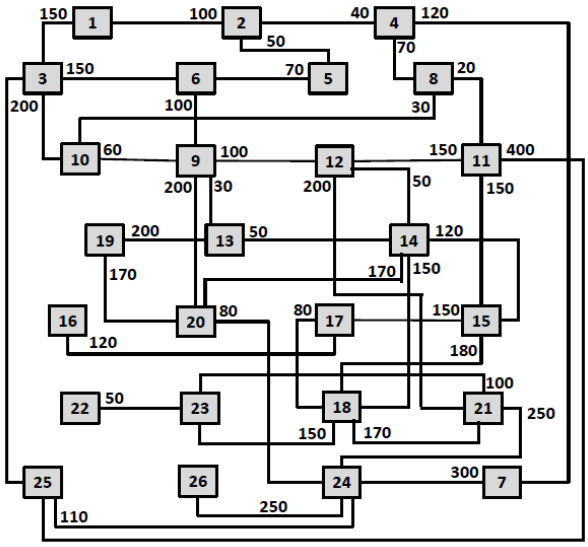


Figure (4.10) The network topology used to simulate various multipath algorithms

modulation format ratio for path 2 (M_2) is 2, it is equivalent to 2 (u_2/M_2) sub-carriers. Similarly, path 3 has 6 available sub-carriers which is equivalent to 3 (u_3/M_3) sub-carriers. Since none of the paths have more than 3 sub-carriers, the single path algorithm would fail here. For the multipath algorithm proposed in [58], 3 paths and 10 sub-carriers are required to satisfy the demand. Note that the demand is 4 sub-carriers in terms of the highest modulation rate format. However, as the modulation format rate for the second ($M_2 = 2$) and the third path is lower ($M_3 = 2$), the number of required sub-carriers is more. Moreover, guard-carrier needs to be added in each path to calculate total number of sub-carriers required. Hence, when applying the multipath algorithm proposed in [58], for the first (shortest) path 1, $u_1/M_1(= 1) < D_m(= 4)$, these sub-carriers are reserved. The remaining demand to be satisfied will be $D_m - u_1/M_1 = 3$. In the second shortest path 2, $uU_2/M_2 = 2$ and it still does not fulfill the remaining demand. Hence, $u_2(= 4)$ sub-carriers are reserved and the algorithm moves on to the third shortest path 3. Now, the remaining demand is $(D_m - u_1/M_1 - u_2/M_2 = 1)$. For path 3, the demand satisfying capacity is $u_3/M_3 = 3$, which is greater than the remaining demand, hence 3 paths ($n = 3$) are required. The number of sub-carriers to be reserved along the third path only corresponds

to the remaining demand ($D_m - u_1/M_1 - u_2/M_2) = 2$). As a result, the full capacity for of path 1 and path 2, plus part of path3 are reserved to satisfy the request. In specific, for the path 1, it will be u_1 + *guard-carrier* = 2; for path 2, it will be u_2 + *guard-carrier* = 5; and for the path 3, it will be the remaining demand + *guard-carrier* = 3. Hence, the total sub-carriers required is 10 when applying the multipath algorithm in [58].

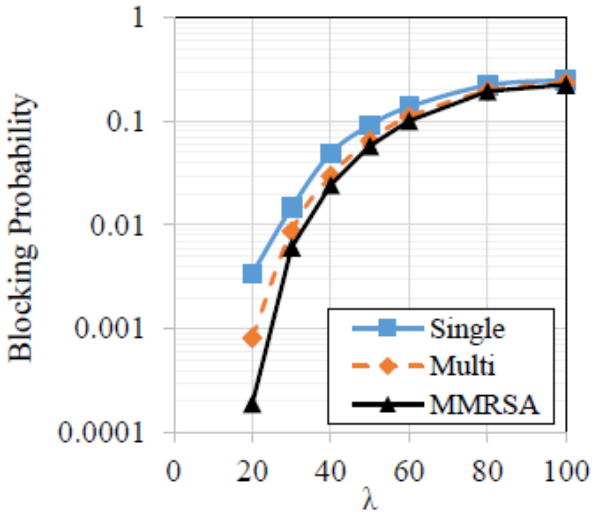
In MMRSA algorithm, the number of shortest paths to satisfy the demand (progressively calculating from the shortest path) is same as the multipath algorithm ($n = 3$). However, in the next step, path 1 and 2 (except path 3) are arranged in descending order of their capacity (u/M). Hence, the order will be path 2 ($u_2/M_2 = 2$) > path 1 (u_1/M_1). Now, different path combinations are compared. The first combination is path 3 (n^{th} or 3^{rd}) which is not sufficient ($u_3/M_3(= 3) < D_m(4)$). The next combination is last path (n^{th}) + first path in the order (i.e., path 2). This combination requires 10 sub-carriers (u_2 + *guard-carrier* = $5, u_3/M_3(D_m - u_2/M_2)$ + *guard-carrier* = 5). The last (n^{th}) combination is path 3 + path 2 + path 1. This combination also requires 10 sub-carriers. Since combination 2 requires only 2 paths, this combination is selected. Note that the MMRSA algorithm also considered the combination from the multipath algorithm in [58] (i.e., case 3 in MMRSA algorithm).

Table (4.1) Parameters used in simulation of MMRSA

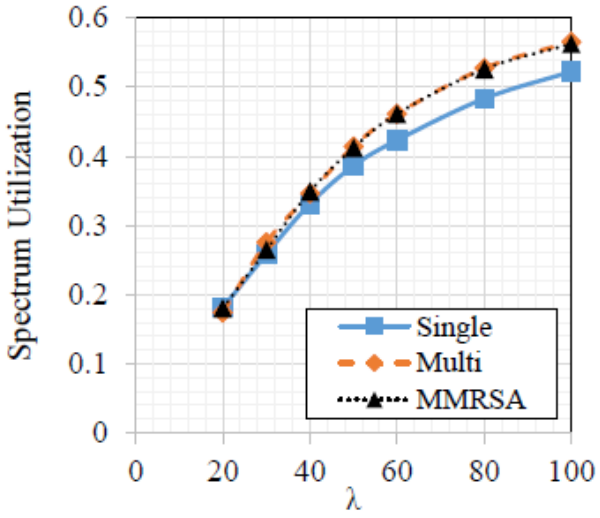
Different Cases	Demand (Gbps)	No. of guard carriers
Case 1 (Base)	800-1000	1
Case 2 (Larger GC)	800-1000	2
Case 3 (Higher demand)	1700-1900	1

4.4.3 Performance Evaluation of MMRSA

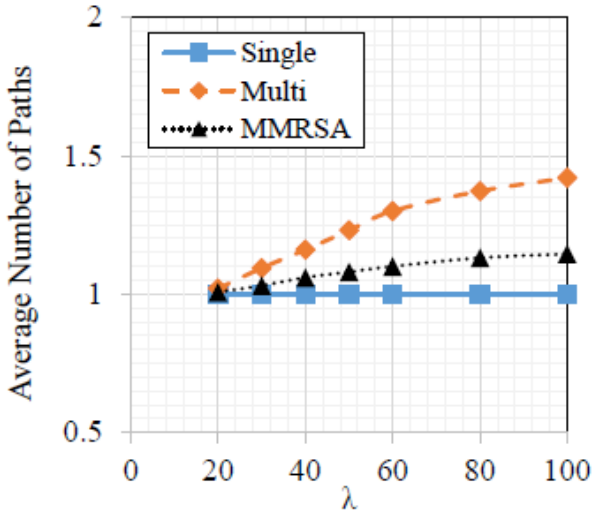
A network with 26 nodes and 41 links as shown in Figure 4.10 is used to compare the performance of the single path, multipath algorithm proposed in [58] with the proposed multipath algorithms. Each link is assumed to have 126 sub-carriers while each sub-carrier is of 12.5 GHz. Three modulation formats - 64 QAM, 16 QAM and QPSK are considered



(a) λ vs. Blocking Probability.

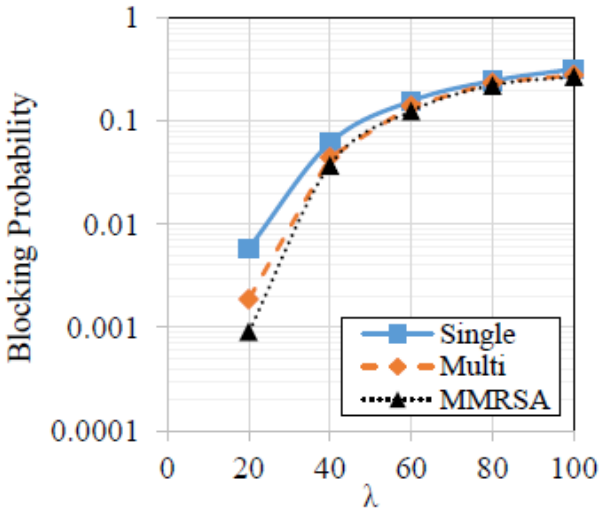


(b) λ vs. Spectrum Usage.

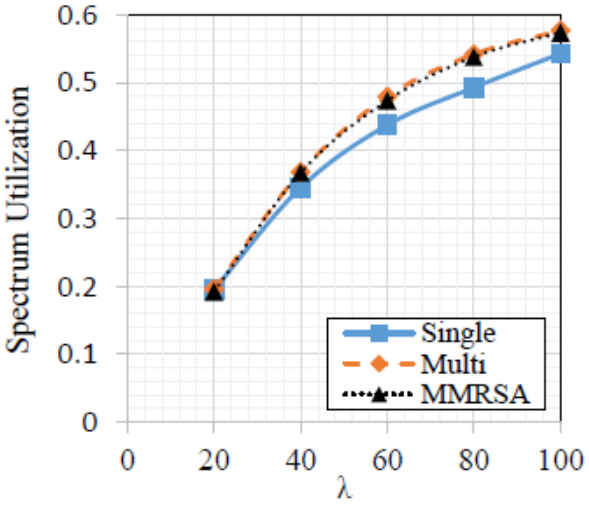


(c) λ vs. Average Number of Paths.

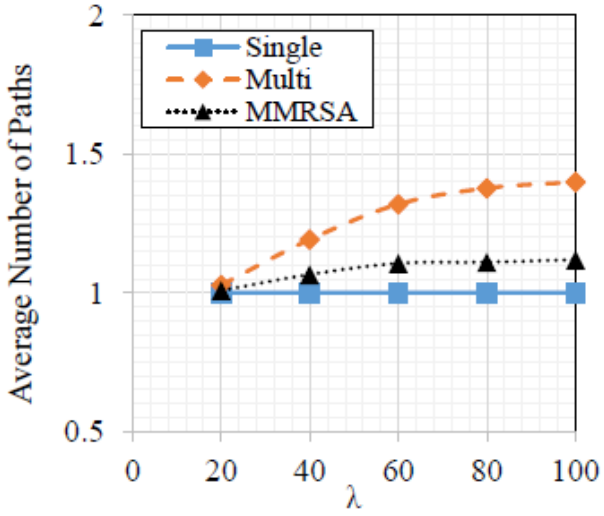
Figure (4.11) Comparison of λ vs. different parameters when GC = 1 and low demand [Case 1]



(a) λ vs. Blocking Probability.



(b) λ vs. Spectrum Usage.



(c) λ vs. Average Number of Paths.

Figure (4.12) Comparison of λ vs. different parameters when GC = 1 and high demand [Case 2]

and assumed to have a modulation format rate of 8 bit/sec, 4 bit/sec and 2 bits/sec resp. and transmission distances of 400, 800, 1600 units resp. The demand with random source and destination is generated uniformly using Poisson's distribution with a mean rate of λ. Different cases that are investigated are summarized in Table 4.1.

Figure 4.11(a) compares the performance of these algorithms in terms of the blocking probability. It can be seen from the figure that the blocking probability increases with the increase in λ while the blocking probability for the MMRSA algorithm is lower than the other two algorithms. However, the blocking probabilities for all the three algorithms are close to each other when the traffic is very high. The performance of λ vs. Spectrum usage is shown in Figure 4.11(b). Spectrum usage is the ratio of number of occupied sub-carriers (including guard carriers) and the total number of sub-carriers across all the links. From the figure, it can be seen that the single path algorithm has the least spectrum usage among the three algorithms. This may be because the single path algorithm has higher blocking probability which means that lesser number of demands is being satisfied and hence results in lower spectrum usage. Also note that the spectrum usage for the MMRSA algorithm is almost the same as that of the multipath algorithm. This further suggests that the spectrum is utilized more efficiently in the MMRSA algorithm as compared to the multipath algorithm. Figure 4.11(c) shows that the MMRSA algorithm requires less number of paths per connection as compared to the multipath algorithm. This can also be understood from the example discussed previously where the MMRSA algorithm utilized lower number of paths as compared to the multipath algorithm. Having lower number of paths may be beneficial as it will avoid splitting the resources into more number of paths and then later combining them. In addition, lower number of paths also decrease the number of GC required in the MMRSA algorithm. Hence, though the multipath and MMRSA algorithms have the same spectrum usage, the MMRSA algorithm utilized it more efficiently be actually assigning it to the connection demands, rather than to the GC. Figure 4.11(c) also suggests that the number of paths per connection is close to 1 when there are less number of connections. When there are less number of connections, a single path is sufficient. But as the number

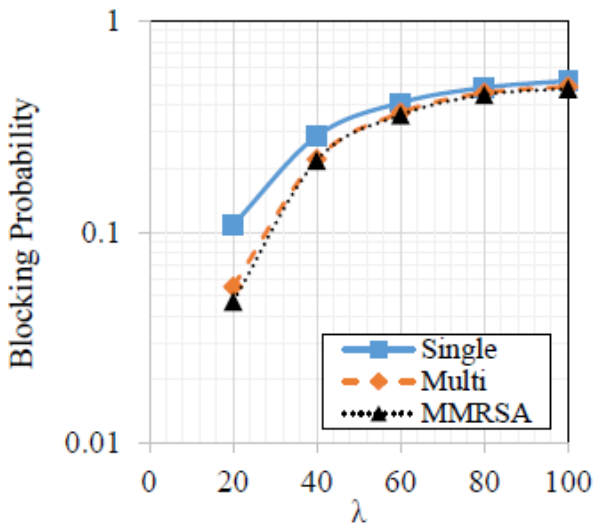
of connections are increased and the spectrum is more fragmented, the MMRSA algorithm utilized multiple paths to yield lower blocking probability.

Figure 4.12(a) shows the blocking probability for case 2. In this case, the size of GC is increased to 2 (from 1 in case 1). From Figure 4.12(a), it can be seen that the blocking probability is higher for all the three algorithms for case 2 than in case 1 as the guard-carriers are expected to consume more spectral resources. However, the MMRSA algorithm still has the lowest blocking probability among the three algorithms. Hence, the MMRSA algorithm outperforms the other two in terms of blocking probability even when the size of the guard carrier is increased. This is also reflected from higher spectrum usage in Figure 4.12(b) for the MMRSA algorithm as compared to the base case. The higher spectrum usage means more spectrum is being utilized which in turn signifies more connection are being accepted. Similar to the base case, the MMRSA algorithm is observed to require less number of paths as seen in Figure 4.12(c) in comparison to the multipath algorithm in case 2 as well.

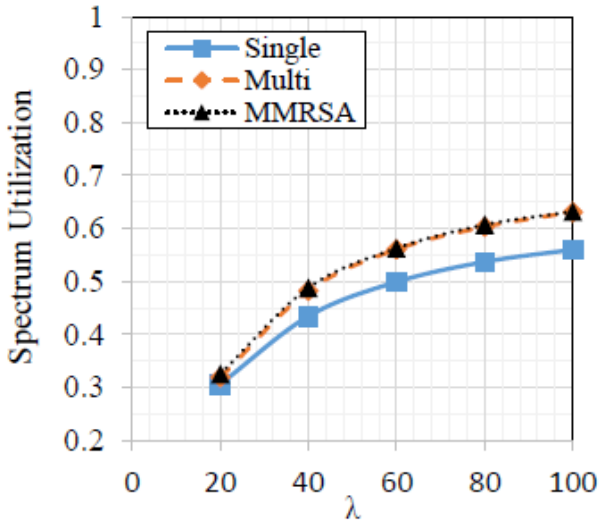
We also measured the performance of these algorithms with higher demands as shown in Figure 4.13. From Figure 4.13(a) and 4.13(b), it can be seen that with the increase in demand size, the blocking probability and the spectrum usage also increases. It can also be seen that the MMRSA algorithm shows slightly better performance than multipath algorithm and requires lower average number of paths as seen in Figure 4.13(c). Note that as the demand increases (similar to increasing the number of connections), more spectrum is occupied and this may cause the algorithms to have closer performance. From the different cases investigated, it can be noted that the MMRSA algorithm performs better as compared to the multipath and single-path algorithms, though the improvement reduces as the demands or number of connections are increased.

4.5 Chapter Summary

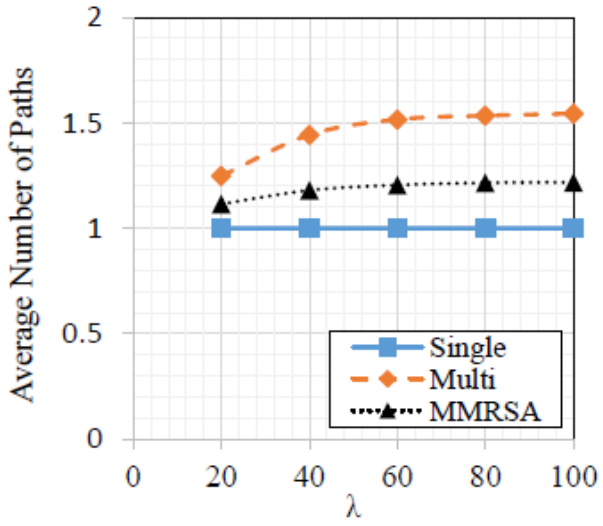
In this chapter, we explained the issue of spectral fragmentation in EONs. We also presented various proactive and reactive defragmentation schemes and also, presented in detail an efficient heuristic algorithm that defragments spectral fragments using Independent sets



(a) λ vs. Blocking Probability.



(b) λ vs. Spectrum Usage.



(c) λ vs. Average Number of Paths.

Figure (4.13) Comparison of λ vs. different parameters when GC = 1 and high demand [Case 3]

with the objective of lowering the sub-carrier index in any fiber. Simulation results confirm that the proposed heuristic can effectively lower the MSI used in any fiber while reducing blocking probability when compared to the dynamic RSA algorithm with no defragmentation. While spectrum fragmentation can be managed, multipath algorithm may be used to cope with the fragmented spectral resources. For this, we presented a novel multipath algorithm with distance-adaptive modulation format in EONs that utilizes minimum SPs called Modulation-aware Multipath Routing Algorithm (MMRSA). The proposed MMRSA algorithm calculates the multiple SPs from source to destination and then uses the combination of SPs that minimizes the total number of paths used. By ordering the SPs in terms of their demand satisfying capacity, the MMRSA algorithm decreases the number of combination to be exposed to n combination. The simulation results showed that the MMRSA algorithm results in lower blocking probability and higher spectrum usage and also, requires less number of paths as compared to the multipath and a single path algorithm for different guard-carrier (GC) size.

PART 5

SPECTRUM ALLOCATION IN ELASTIC OPTICAL NETWORKS USING TRAFFIC PREDICTION

In the previous chapters, we have already discussed the RSA problem in EONs. For the case with static traffic, the RSA problem is shown to be NP-Complete [6]. The RSA problem with dynamic traffic can also be very challenging due to the random traffic arrival/departure time. When the traffic demand fluctuates over time, a straightforward solution is to reconfigure the spectrum assignment periodically (or on demand) by doing the RSA process from scratch based on the traffic demands. This approach, however may require shifting many (if not all) existing/operating sub-carriers, reconfiguring network elements (e.g. transponders and receivers), and interrupting the live connections, which can be very expensive for the network providers. In this chapter, we study how to assign sub-carriers in EONs or Flex-Grid networks with traffic demands fluctuating over time while minimizing the network reconfiguration. We model the time-varying traffic allocation problem using the Integer Linear Programming (ILP) formulations and develop heuristics based on the Interference Graph technique to effectively assign sub-carriers while minimizing the network reconfiguration cost.

5.1 Related Work in Spectrum Allocation

In the literature, some work has been done to assign sub-carriers to connection requests in networks with dynamic traffic demands. A policy that allocates sub-carriers to requests whose rate fluctuate with time was proposed in [53]. In [64], an adaptive spectrum allocation algorithm for time-dependent traffic was investigated to share neighboring sub-carriers based on the match factor between two neighboring connections. However, in real network scenarios, this scheme may not be able to find many neighboring connections with a good match. The authors in [65] proposed a new method of sharing the sub-carriers between

the primary traffic and the secondary traffic. The primary traffic is the time-dependent traffic that fluctuates with time whereas the secondary traffic is the dynamic traffic that comes and leaves randomly. Hence, the authors in this paper proposed that the secondary traffic can utilize the over-allocated sub-carriers of the primary traffic when they are not in use by the primary traffic. The authors also used interference factor to avoid interference between primary and secondary traffic. The authors in [66] studied the spectrum sharing algorithm between neighboring connections where the spectrum gap between two neighboring connections can be reduced depending on the sub-carriers required by the neighboring connections. This yields a lower blocking probability and higher spectrum efficiency than no sharing methods. Spectrum expansion, contraction and reallocation policies were studied in [67] which try to reduce the blocking probability by introducing cooperative "neighbor avoidance" mechanisms. It was also shown that if the spectrum is allowed to move freely without imposing restrictions on their reference frequency, it can yield maximum benefits. The work in [68] studied a Multi-Hour Routing and Spectrum Allocation (MHRSA) problem and the authors in [69] explored the periodic behavior of internet traffic in WDM networks with traffic grooming techniques. However, network reconfiguration and traffic prediction in EONs or Flex-Grid networks with dynamic time-varying traffic have not been considered in the existing work. In this chapter, we present on how to effectively accommodate time-varying traffic with minimum network reconfiguration and blocking probability in EONs or Flex-Grid networks [70], [71], [72].

The remainder of this chapter is organized as follows. Section 5.2 presents the time-varying spectrum allocation problem in EONs while Section 5.4 formulates the time-varying traffic allocation in EONs or Flex-Grid Networks as an Integer Linear Programming problem. Section 5.5 presents four heuristics schemes to allocate sub-carriers for the time-varying traffic. Section 5.6 shows the performance evaluation of the proposed schemes and Section 5.7 presents the summary of this chapter.

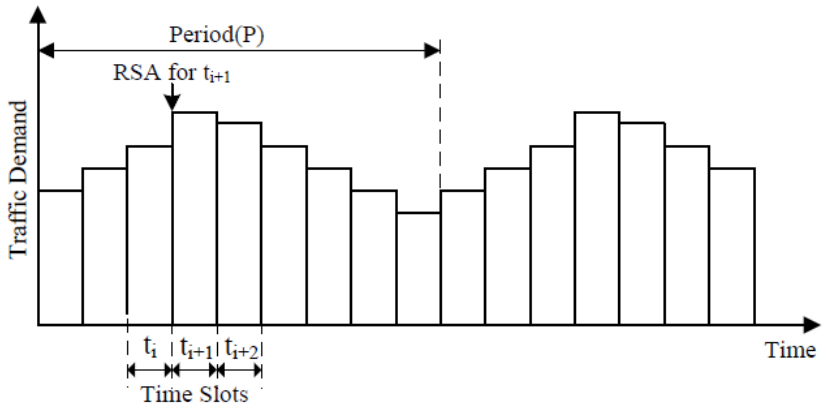


Figure (5.1) An example of time-varying traffic.

5.2 Time-varying traffic in EONs

It has been observed that although the real-world data traffic may fluctuate randomly in a very short period of time (e.g. within one hour), such a change follows a similar (or periodic) pattern over a long period of time (e.g., a day or a month) [73], [74]. More specifically, the Internet traffic exhibits a similar pattern every 24 hours [75]. This is largely due to the habitual usage of the network users. The similar or periodic traffic pattern over a long period has been referred to as *time-varying traffic* [69]. For *time-varying traffic*, we can assume that the traffic pattern repeats after a certain time period (P). For example in Figure 5.1, the traffic pattern repeats in a time period of P, which can be divided into n fixed-length time slots (e.g. one hour time slot). To handle *time-varying traffic*, we assume a slotted control plane in EONs where the traffic demand between each node-pair is almost constant within a time slot and the highest intensity of the flow is used as the traffic demand in that time slot. The significant traffic change only happens at the boundary of the time slot. We further assume that the traffic requests for the next time slot say t_{i+1} are given before the end of the current time slot t_i. In addition, at time slot t_i, the traffic requests for the time slot t_{i+2} are not precisely known. The process of RSA at time slot t_{i+1} will occur at the end of time slot t_i as indicated in Figure 5.1.

Consider a network as *G(V, E, N)* where V represents the set of nodes, E represents

the set of fiber links between nodes in V, and N is the maximum number of sub-carriers in each fiber. Then, we can define the Time-varying Spectrum Allocation (TSA) problem in EONs as follows -

Time-varying Spectrum Allocation (TSA) problem: Given a set of spectrum path assignments at time slot $t-1$, $D_{t-1} = \{f_{t-1}\}$ on a network $G(V, E, N)$, and a set of demands for the next time slot $D_t = \{f_t\}$, is it possible to accommodate all the flows in demand D_t with minimum interruptions to the existing traffic flows?

Theorem 1. *The time-varying spectrum allocation problem (TSA) in EONs is NP-hard.*

Proof. At any time slot t, the TSA problem has to allocate sub-carriers to all the SPs in demand D_t. Hence, the TSA problem consists of an instance of the static spectrum allocation (SSA) problem, which can be defined as - Given a network $G(V, E, N)$ and demands between node-pairs, is it possible to establish SPs for all the demands while satisfying the *spectrum-continuity* and *sub-carrier consecutiveness* constraints? The SSA problem has shown to be NP-Complete [6] via its connection with the static lightpath establishment (SLE) problem in wavelength routed networks. The SLE problem is shown to be NP-Complete [76] which employs a reduction to (and from) n-graph colorability theorem. Since SSA problem is known to be NP-Complete [76], the TSA problem which consists of an instance of the SSA problem in EONs or Flex-Grid networks is at least as hard as the SSA problem. □

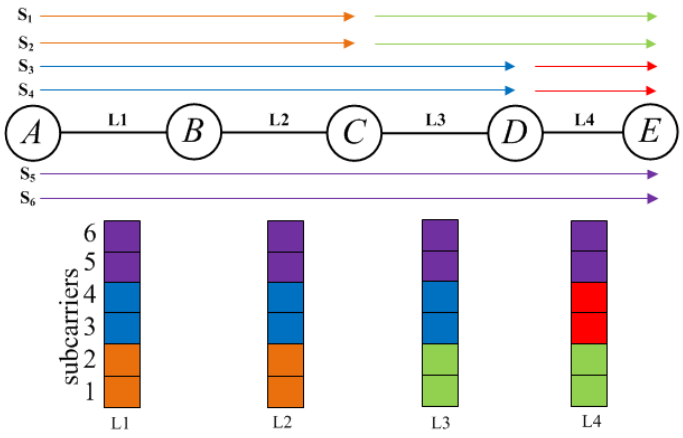


Figure (5.2) Existing spectrum assignment at time slot t_i.

Table (5.1) Traffic demand at time slots t_i, t_{i+1} and t_{i+2}.

Node-pair	t_i	t_{i+1}	t_{i+2}	Δ_{i+1}	Δ_{i+2}
(A, C)	2	2	3	0	1
(C, E)	2	2	3	0	1
(A, D)	2	3	4	1	1
(D, E)	2	3	3	1	0
(A, E)	2	2	2	0	0

5.3 Spectrum Allocation to Time-varying traffic in EONs

Figure 5.2 shows a linear network with the existing spectrum assignments between node-pairs at time slot t_i. The expected traffic demands between the node-pairs at time slots t_{i+1} and the predicted traffic demands at time slot t_{i+2} are given in Table 5.1. If we denote the traffic demand of a given node-pair at time slot t_i by $n^d_{t_i}$, then we can use the following equation to calculate the traffic change between a node-pair from time slot t_i to time slot t_{i+1}.

$$\Delta_{i+1} = n^d_{t_{i+1}} - n^d_{t_i}$$

The value of Δ_{i+1} can be either negative, positive or zero. The negative value of Δ_{i+1} indicates the decrease in traffic demand. This may create vacant or unused sub-carriers but may not require reconfiguration of network elements such as transponders and receivers. The positive value indicates the increase in traffic demand. If the sub-carriers contiguous to this demand are occupied, sub-carrier reallocation and network reconfiguration have to be carried out resulting in possible interruptions to some live connections. Moreover, if there are not enough empty sub-carriers to relocate this demand, such demand could be either dropped or blocked. When there are vacant sub-carriers adjacent to this demand, then we may reduce the reconfiguration overhead by simply stretching the original *SP* instead of reallocating this *SP* to a new set of sub-carriers.

For example, in Figure 5.2, node-pair (A, D) uses sub-carriers S_3 and S_4 at time slot t_i. At time slot t_{i+1}, if the sub-carrier S_5 were available to expand for node-pair (A, D) to 3 sub-carriers, then node-pair (A, D) would not have to be assigned with a different set of sub-carriers. In Figure 5.2, the sub-carrier S_5 is occupied by the *SP* between node-pair (A, E). Hence, in order to free up the sub-carrier S_5 for node-pair (A, D), the *SP* for node-pair (A, E) may have to be moved from sub-carriers S_5, S_6 to the different set of sub-carriers. If there are not enough sub-carriers to relocate *SP* for node-pair (A, E), then the connection request of 3 sub-carriers between node-pair (A, D) may have to be blocked.

Reallocation of connections from one set of sub-carriers to a different set of sub-carriers can be very costly to network providers as it involves re-tuning optical lasers and may also require additional resources such as transponders or bandwidth-variable optical cross connects. Compared to the case of reallocating an *SP* to a totally different set of sub-carriers, the reconfiguration cost for stretching the sub-carriers to vacant adjacent sub-carriers may be negligible since the connection of the traffic demand is not fully interrupted and it may not require additional resources [77]. Similarly, when the traffic demand of a node-pair decreases, the reconfiguration involves turning-off operations of network elements, whose cost may also be negligible.

5.4 ILP model for the time-varying traffic problem

In this section, we formulate the time-varying traffic allocation problem in EONs or Flex-Grid networks for one time period as an Integer Linear Programming (ILP) problem using the *path-link* approach [25].

5.4.1 Notation and variables

The following notations and variables are required in the proposed mathematical ILP model.

$G(V, E, N)$	The physical topology of the network where V is the number of nodes, E is the set of fiber links between nodes in V and N is the number of sub-carriers in each link with sub-carrier set $S = \{S_1, S_2, ...S_N\}$ and $N = \|S\|$;
t	A specific time slot $t \in \{1, 2, ..., T\}$, where T is the number of time slots in a period,
D_t	The set of demands at time slot t; a specific demand $d \in D_t$,
$n_t^d \in \mathbb{Z}_+$	The number of sub-carriers requested by demand d at time slot t,
P	Set of all paths, and $P_e \subseteq P$ identifies all paths that go through link e.
P_d	Predetermined set of candidate paths for demand d,
x_{spdt}	Equal to 1 if sub-carrier s is assigned as the lowest indexed sub-carrier in path p for demand d in time slot t and equal to 0 otherwise,
y_{spdt}	Equal to 1 if sub-carrier s is assigned in path p for demand d at time slot t and equal to 0 otherwise,
C_{pdt}	Equal to 1 if the spectrum assignment at path p changes from time slot $t-1$ to t for demand d, and equal to 0 otherwise,
MS	The maximum index of the sub-carriers allocated among all the fibers in the network.

5.4.2 The objective function

The objective of the proposed ILP model can be to minimize the maximum index of the sub-carriers allocated among all the fiber links or, to minimize the number of network reconfiguration in one time period which are specified in Eq. (5.1) and Eq. (5.2) resp.

$$min \quad MS, \tag{5.1}$$

$$min \sum_{d \in D_t} \sum_{p \in P_d} \sum_{t \in T} C_{pdt}, \tag{5.2}$$

5.4.3 Constraints

Following are the constraints in the proposed ILP model.

$$\sum_{p \in P_d} \sum_{s \in S} x_{spdt} = 1, \quad \forall d \in D_t, \forall t \in \{1, 2, ..., T\} \tag{5.3}$$

$$\begin{aligned} x_{s_i pdt} - y_{s_j pdt} \leq 0, \quad & \forall d \in D_t, \forall p \in P_d, \forall s_i, s_j \in S, \\ & t \in \{1, 2, ..., T\} \\ & where\ i = 1 \ldots |S| - n_t^d + 1, \\ & and\ j = i, \ldots, i + n_t^d - 1 \end{aligned} \tag{5.4}$$

$$\begin{aligned} x_{s_i pdt} = 0, \quad & \forall d \in D_t, \forall p \in P_d, \forall s_i \in S, \\ & where\ i = |S| - n_t^d + 2, \ldots, |S| \end{aligned} \tag{5.5}$$

$$\sum_{p \in P_d} \sum_{s \in S} y_{spdt} = n_t^d, \quad \forall d \in D_t, \forall t \in \{1, 2, ..., T\} \tag{5.6}$$

$$\sum_{p \in P_e} y_{spdt} \leq 1, \quad \forall e \in E, \forall s \in S, t \in \{1, 2, ..., T\} \tag{5.7}$$

$$MS \geq i \cdot x_{s_i pdt} + n_t^d, \quad \forall s_i \in S, \forall p \in P_d, \forall d \in D_t, \forall t \in \{1, 2, ..., T\} \tag{5.8}$$

$$C_{pdt} = \frac{1}{2} \sum_{i=1}^{N} x_{s_i pdt} \oplus x_{s_i pdt-1} \quad \forall p \in P_d, \forall d \in D_t, \forall t \in \{1, 2, ..., T\} \tag{5.9}$$

Eq. (5.3) ensures that every demand d is assigned a path p and a sub-carrier s is selected as the lowest indexed sub-carrier in that path at time slot t. Eq. (5.4) and (5.5) ensure that the demands are assigned with sub-carriers following the *sub-carrier consecutiveness* constraint. Whenever there is a sub-carrier s_i selected as the lowest indexed slot for demand d in path p, the consecutive slots s_j, where $j = i, ..., i + n_d - 1$, should be assigned to this demand. The *spectrum-continuity* constraint is imposed implicitly along the path p since the spectrum assignment is done along the entire path p. Eq. (5.6) ensures that each demand d is assigned with the requested number of sub-carriers. Eq. (5.7) is the sub-carrier occupancy constraint which guarantees that the sub-carrier in each network link can be assigned to at most one demand. Eq. (5.8) is used to obtain the maximum index of the sub-carriers allocated among all the fiber links. Meanwhile, Eq. (5.9) summarizes the changes in spectrum assignment in two consecutive time slots $t-1$ and t along the same path. The XOR operation in Eq. (5.9) counts each change in spectrum assignment twice, first at the sub-carrier where the demand was previously assigned at time slot $t-1$ and then at the sub-carrier where the demand is currently assigned at time slot t. Hence, the result is multiplied by $\frac{1}{2}$. Note that, the XOR operation in Eq. (5.9) is not linear. To make it linear, the XOR operation (i.e. $x_{s_i pdt} \oplus x_{s_i pdt-1}$) in Eq. (5.9) can be replaced with an integer variable z that can be computed using linear equations from Eq. (5.9a) - Eq. (5.9e).

$$z \leq x_{s_i pdt} + x_{s_i pdt-1} \tag{5.9a}$$

$$z \geq x_{s_i pdt} - x_{s_i pdt-1} \tag{5.9b}$$

$$z \geq x_{s_i pdt-1} - x_{s_i pdt} \tag{5.9c}$$

$$z \leq 2 - x_{s_i pdt} - x_{s_i pdt-1} \tag{5.9d}$$

$$0 \leq z \leq 1 \tag{5.9e}$$

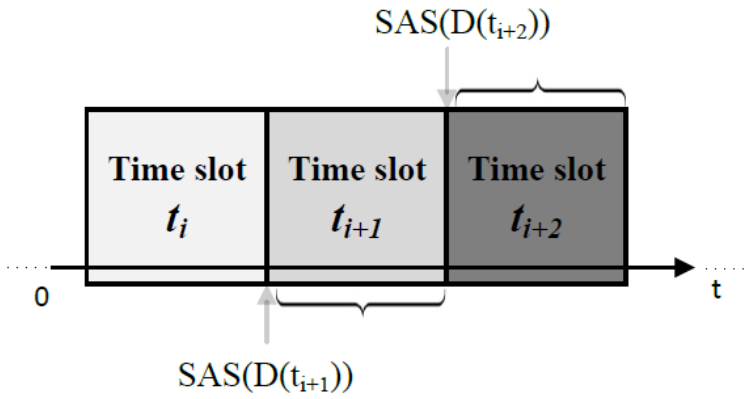


Figure (5.3) Spectrum allocation in SAS.

5.5 Heuristic spectrum allocation algorithms for time-varying traffic

The above ILP model is intractable for large problems due to the high computational complexity. Hence, in this section, we present four fast heuristic algorithms to assign sub-carriers to *time-varying traffic*, namely, Spectrum Allocation from Scratch (SAS), Planning-ahead Spectrum Allocation (PSA), Predictive Planning-ahead Spectrum Allocation (PPSA) and Predictive Spectrum Band Allocation (PSBA). We assume that all the heuristics only accommodate a request with one consecutive spectrum blocks along a single path and otherwise block the request [1]. PSA and PPSA utilize Minimum Reconfiguration spectrum

[1]To avoid potential out-of-order packets arrival on multiple paths and degradation of resource usage, we assume that a request can only be satisfied by a set of consecutive sub-carriers along the same path.

Allocation (MRA) algorithm to minimize the network reconfiguration cost while allocating sub-carriers to SPs. Hence, MRA will be explained first before describing PPA and PPSA. For the proposed algorithms, reallocation of SPs can cause network disruption in EONs [39]. Hence, to minimize disruption, we use a make-before-break scheme. Using this scheme, *SPs* are first established on different set of sub-carriers before releasing the original SP which can minimize the disruption time [39].

5.5.1 Spectrum Allocation from Scratch (SAS)

Spectrum Allocation from Scratch (SAS) adopts a straightforward method to allocate sub-carriers to each SP based on the current traffic demands. At the end of time slot t_i, the traffic demands for next time slot t_{i+1} are given in a set $D(t_{i+1}, s, d)$. SAS only considers the traffic demands at time slot t_{i+1} in the process of spectrum allocation as shown in Figure 5.3. The detailed steps for SAS are given in Algorithm 3. SAS is a naive approach and may result in frequent set-up and tear-down of connections due to the traffic fluctuations.

Algorithm 3 Spectrum Allocation from Scratch (SAS)

Input: $\{D(t_{i+1}, s, d)\}$
Output: $\{SA(t_{i+1}, s, d)\}$

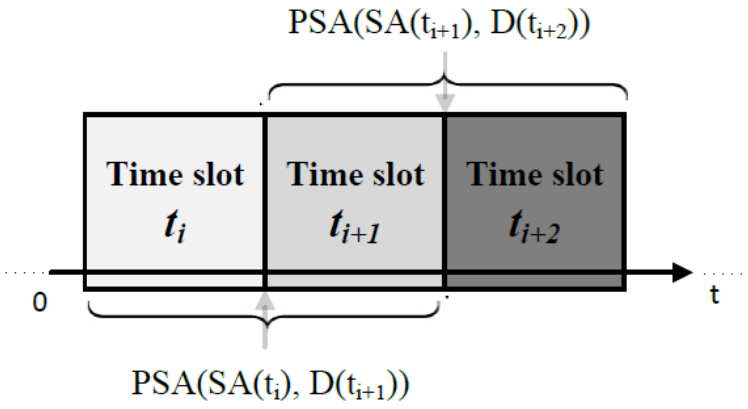
```
Sort the traffic demands in the descending order of sub-carriers requested
while there are requests in D(t_{i+1}, s, d) do
  if the request is unchanged in t_{i+1} then
    Assign them the same sub-carriers as in t_i
  else
    Find the shortest route for (s, d)
    Use First-fit scheme to find D(t_{i+1}, s, d) sub-carriers along the shortest route
    if found available sub-carriers for (s, d) then
      Allocate those sub-carriers to (s, d)
    else
      Block the request for (s, d)
    end if
    Update SA(t_{i+1}, s, d) and D(t_{i+1}, s, d)
  end if
end while
```

5.5.2 Minimum Reconfiguration Spectrum Allocation Algorithm

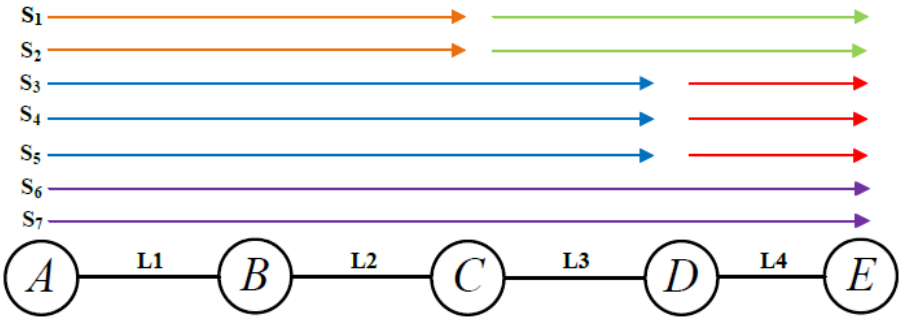
For a *time-varying traffic*, at time slot t_i, the spectrum assignment for demands at t_i is known. The traffic demands at next time slot t_{i+1} are also known while the traffic demands at time slot t_{i+2} can be predicted based on the periodic behavior of Internet traffic. These information can be utilized during the sub-carrier allocation process in order to minimize the network reconfiguration cost. In this section, we now describe the Minimum Reconfiguration spectrum Allocation (MRA) algorithm that will utilize the traffic demands and the predicted traffic demands at future time slots to minimize the total reconfiguration cost over time.

MRA consists of two steps. In Step 1, we transform the spectrum allocation information at current time slot t_i into an auxiliary Interference Graph (IG). To construct the IG, we create a vertex for each SP at the current time slot t_i and connect two vertices (say i and j) if sub-carriers of j have to be reallocated for the accommodation of vertex i in the future time slots. In Step 2, we find the Maximum Independent Set (MIS) in the IG to calculate the configuration cost. Finding the MIS in a graph is a well-known NP-Complete problem. There is a parallel version of algorithm [78] which can achieve $O((logn)^4)$ with a space complexity of $O((n/logn)^3)$. For the simplicity of the implementation, we use the algorithm in [79] to find the MIS in an IG.

The number of nodes in the IG depends on the number of SPs while the number of edges depends on the spectrum demands between node-pairs at future time slots. Since, at the end of the time slots t_i, the spectrum demands at time slot t_{i+1} are known while the spectrum demands at time slot t_{i+2} can be predicted, we propose two heuristics that utilize MRA: Planning-ahead Spectrum allocation (PSA) and Predictive Planning-ahead Spectrum allocation (PPSA).



(a) Spectrum allocation in PSA.



(b) An example of PSA at time slot t_{i+1}.

Figure (5.4) Planning-ahead Spectrum Allocation (PSA).

Planning-ahead Spectrum Allocation (PSA) Planning-ahead Spectrum Allocation (PSA) considers two factors for the spectrum assignment between node-pairs *(s, d)* as shown in Figure 5.4 (a): $SA(t_i, s, d)$ and $D(t_{i+1}, s, d)$ where the former is the spectrum assignment at time slot t_i and the latter is the traffic demand at time slot t_{i+1}.

For example, in Figure 5.2, the node-pairs (A, D), (D, E) and (A, E) are assigned with 2 sub-carriers at time slot t_i. At time slot t_{i+1}, the traffic demands between node pair (A, D) and (D, E) are 3 sub-carriers (i.e. $\Delta_{i+1} = 1$). Hence, to minimize the network reconfiguration cost, PSA may assign node pair (A, E) with sub-carriers S_6 and S_7 as shown in Figure 5.4 (b). This allows the SPs between node pairs (A, D) and (D, E) to utilize the consecutive sub-carrier S_5 during time slot t_{i+1}. Hence, in this scheme, only SP assigned to node-pair

(A, E) has to be reallocated while *SPs* assigned to node-pair (A, D) and (D, E) can operate without interruptions.

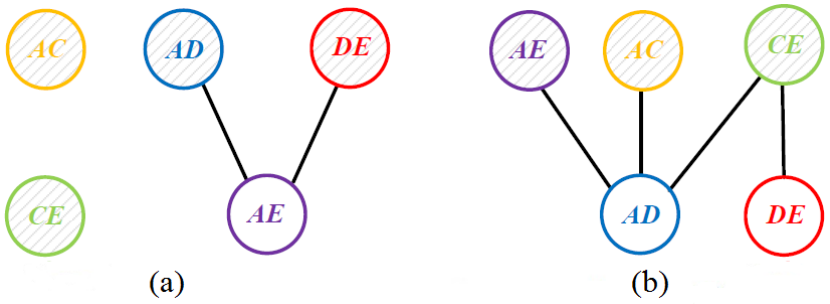


Figure (5.5) PSA interference graph for (a) time slot t_{i+1} (b) time slot t_{i+2}.

For the spectrum assignment given in Figure 5.4 (b), the *IG* of PSA at time slots t_{i+1} and t_{i+2} are shown in Figure 5.5(a) and (b), respectively. There are 5 *SPs* at time slot t_i which are assigned with the sub-carriers as shown in Figure 5.2. Hence, in the IG, there are 5 vertices : *AC*, *AD*, *DE*, *CE* and *AE*. There is an edge between vertices *AD-AE* and *DE-AE* since *SP* for node-pair (A, E) has to be moved to provide sub-carrier S_5 for node-pairs (A, D) and (D, E) to expand at time slot t_{i+1}. With the *IG*, we can then calculate the reconfiguration cost by finding the MIS among the vertexes. The MIS of the *IG* at time slot t_{i+1} consists of shaded nodes *AC*, *AD*, *DE* and *CE* as shown in Figure 5.5(a). Similarly, at time slot t_{i+2}, the size of the MIS is 3. Hence, the reconfiguration cost at time slot t_{i+1} is 5 - 4 = 1 since only one *SP* needs to be interrupted and relocated (i.e. *SP* for node-pair (A, E)). The total reconfiguration cost for PSA at time slots t_{i+1} and t_{i+2} is $(5-4)+(5-3) = 3$.

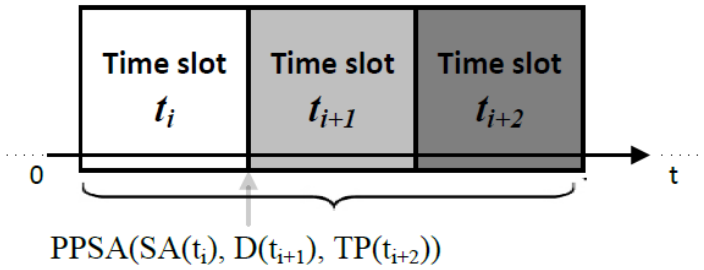
Predictive Planning-ahead Spectrum Allocation (PPSA) Predictive Planning-ahead Spectrum Allocation (PPSA) takes advantage of the periodic behavior of the Internet traffic demands. PPSA considers three factors during the spectrum assignment process: the spectrum assignments at time slot t_i given by $SA(t_i, s, d)$, the traffic demands at time slot t_{i+1} given by $D(t_{i+1}, s, d)$ and the predicted traffic demands at time slot t_{i+2} given by $TP(t_{i+2}, s, d)$.

Algorithm 4 Planning-ahead Spectrum Allocation (PSA)

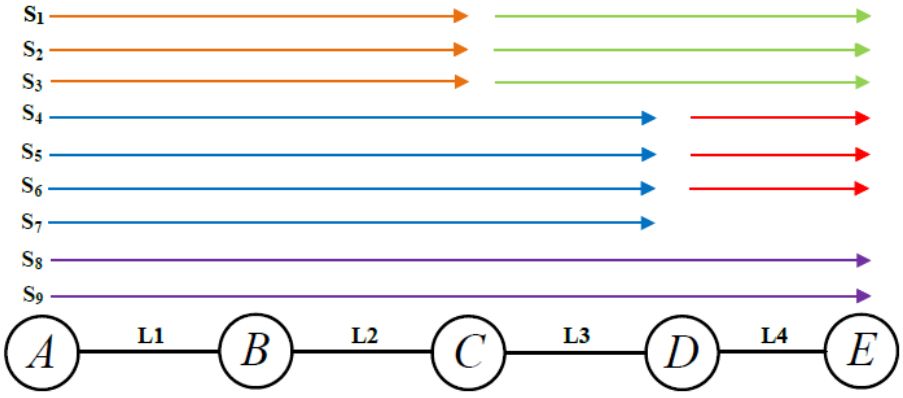
Input: $\{SA(t_i, s, d), D(t_{i+1}, s, d)\}$
Output: $\{SA(t_{i+1}, s, d)\}$

$S_{aux} \longleftarrow \emptyset$
$R \longleftarrow \emptyset$
$P_{aux} \leftarrow 0$, $P \leftarrow 0$
Sort the traffic demands in the descending order of sub-carriers requested
while there are requests in $D(t_{i+1}, s, d)$ **do**
  P = shortest path for *(s, d)*
  if auxiliary graph for $P \notin S_{aux}$ **then**
    Create auxiliary graph for SPs in P using $\{SA(t_i, s, d), D(t_{i+1}, s, d)\}$ as P_{aux}
    $S_{aux} \longleftarrow S_{aux} \cup \{P_{aux}\}$
  else
    P_{aux} = auxiliary graph retrieved for P from S_{aux},
    Record the SPs to be reallocated in P_{aux} and store it in Set R
  end if
  $relocation \leftarrow true$
  while R is not empty **do**
    Relocate SPs in R
    if SP relocation fails **then**
      Revert the spectrum assignment for SP in R
      $relocation \leftarrow false$
      break;
    end if
  end while
  if $relocation ==$ true **then**
    Stretch the sub-carriers of *(s, d)* to $D(t_{i+1}, s, d)$ sub-carriers
  end if
  if $relocation ==$ false **then**
    Allocate $D(t_{i+1}, s, d)$ sub-carriers to *(s, d)* using *First-fit* scheme
    Block the request of *(s, d)* if not enough sub-carriers
  end if
  Update $SA(t_{i+1}, s, d)$ and $D(t_{i+1}, s, d)$
end while

For example in Figure 5.6(a), the spectrum assignments for time slot t_{i+1} happen at the end of time slot t_i where $SA(t_i, s, d)$ and $D(t_{i+1}, s, d)$ are known while $TP(t_{i+2}, s, d)$ can be predicted using the Internet traffic history. In PSA as shown in Figure 5.4, one needs to further reconfigure the network at the end of time slot t_{i+1} to accommodate 4 sub-carriers for node-pair (A, D). In contrast, if the projected traffic demand at time slot t_{i+2} is also considered, the proposed PPSA can shift the SP between node-pair (A, E) to sub-carriers S_8 and S_9 at time slot t_{i+1} while reserving sub-carrier S_3 for node-pairs (A, C) and (C, E) as shown in Fig 5.6(b). Similarly, sub-carrier S_7 can be virtually reserved for node-pair (A, D) to accommodate the increase in traffic demands at time slot t_{i+2}, which can further reduce the overall network reconfiguration cost.



(a) Spectrum allocation in PPSA



(b) An example of PPSA at time slot t_{i+1}

Figure (5.6) Predictive Planning-ahead Spectrum Allocation (PPSA).

The *IG* of PPSA at time slot t_{i+1} considers the predicted traffic demands at time slot t_{i+2} in addition to the known traffic demands at time slot t_{i+1} and the spectrum assignments at time slot t_i. For example, the *IG* at time slot t_{i+1} is shown in Figure 5.7 when PPSA is employed, provided that the traffic demands at time slot t_{i+2} are predicted as given in Table 5.1. The total reconfiguration cost for PPSA at time slots t_{i+1} and t_{i+2} is $5 - 3 = 2$.

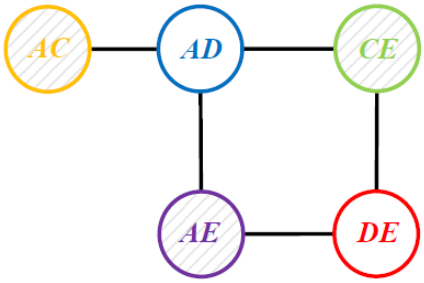


Figure (5.7) PPSA interference graph at time slot t_{i+1}.

Table (5.2) Traffic demand at time slots t_i, t_{i+1} and t_{i+2}.

Node-pair	t_i	t_{i+1}	t_{i+2}	Δ_{i+1}	Δ_{i+2}
(A, C)	2	3	2	1	-1
(C, E)	2	3	2	1	-1
(A, E)	3	2	3	-1	1

5.5.3 Predictive Spectrum Band Allocation (PSBA)

PPSA may reserve a number of sub-carriers to be used by *SPs* for the expansion at the future time slots. These reserved sub-carriers may increase the maximum used sub-carrier index and may lead to poor spectrum utilization. Moreover, if those reserved sub-carriers are not filled in the consecutive time slots, it could lead to spectral gaps in between *SPs*. This could further aggravate the spectrum defragmentation [9]. However, in any time slot,

Algorithm 5 Predictive Planning-ahead Spectrum Allocation (PPSA)

Input: $\{SA(t_i, s, d)\}$, $D\{(t_{i+1}, s, d)\}$ and $\{TP(t_{i+2}, s, d)\}$
Output: $\{SA(t_{i+1}, s, d)\}$

$S_{aux} \longleftarrow \emptyset$
$R \longleftarrow \emptyset$
$P_{aux} \leftarrow 0$, $P \leftarrow 0$
Sort the traffic demands in the descending order of sub-carriers requested
while there are requests in $D(t_{i+1}, s, d)$ **do**
  P = shortest path for *(s, d)*
  if auxiliary graph for $P \notin S_{aux}$ **then**
    Create auxiliary graph for SPs in P using $\{SA(t_i, s, d)\}$, $D\{(t_{i+1}, s, d)\}$ and $\{TP(t_{i+2}, s, d)\}$ as P_{aux}
    $S_{aux} \longleftarrow S_{aux} \cup \{P_{aux}\}$
  else
    P_{aux} = auxiliary graph retrieved for P from S_{aux},
    Record the SPs to be reallocated in Set R
  end if
  $relocation \leftarrow true$
  while R is not empty **do**
    Relocate SPs in R
    if SP relocation fails **then**
      Revert the spectrum assignment for SP in R
      $relocation \leftarrow false$
      break;
    end if
  end while
  if $relocation ==$ true **then**
    Stretch the sub-carriers of *(s, d)* to $max(TP(t_{i+2}, s, d), D(t_{i+1}, s, d))$ sub-carriers
  end if
  if $relocation ==$ false **then**
    Allocate $D(t_{i+1}, s, d)$ sub-carriers to *(s, d)* using *First-fit* scheme
    Block the request of *(s, d)* if not enough sub-carriers
  end if
  Update $SA(t_{i+1}, s, d)$ and $D(t_{i+1}, s, d)$
end while

if one SP is shrinking, the other expanding SP can use those empty sub-carriers left vacant by the shrinking SP (if the two SPs have some links in common). This is like sharing sub-carriers by two neighboring SPs that expand and shrink at the alternate time slots. This leads us to a Predictive Spectrum Band Allocation (PSBA) where an SP with increasing (or decreasing) spectrum demand is grouped (banded) with other SP with decreasing (or increasing) demand along their common path provided that at any time slot, their sub-carrier demand does not exceed the sub-carriers assigned to the band. Here, we define a band factor (B_f) which shows the likelihood of two SPs to form a band. Based on the B_f, two SPs are banded together to form Spectrum Band (SB). The B_f between two neighboring spectrum paths, SP_α and SP_β is given by the following ratio -

$$B_f = \frac{I_{SP_\alpha}}{D_{SP_\beta}} \tag{5.10}$$

where I_{SP_α} is the increase in sub-carrier demand for SP_α and D_{SP_β} is the decrease in sub-carrier demand for SP_β. If $B_f = 1$, then the decrease in spectrum demand of one SP is compensated by the increase in spectrum demand of another SP. If $B_f > 1$, then extra sub-carriers have to be assigned to the band since the increase in spectrum demand of one SP is greater than the decrease in spectrum demand of another SP. However, $B_f < 1$, then the sub-carriers left vacant by the decrease in the sub-carrier demand of one SP is not filled by the increase in sub-carrier demand of another SP. Hence, vacant slots will be created in between SPs, i.e., in between the band. This will create spectrum fragments and reduce spectrum efficiency. Hence, when SPs are banded, preference is given to those neighboring SPs who have a $B_f = 1$, and then to those with $B_f > 1$ and finally to those with $B_f < 1$.

An increasing SP_α can match with a decreasing SP_β based on B_f to create a spectrum band ($SB_{\alpha\beta}$). The traffic demand for $SB_{\alpha\beta}$ at time slot t_{i+1} is $D_{\alpha\beta} = max(D(\alpha_{t_{i+1}}) + D(\beta_{t_{i+1}}))$ and the predicted traffic demand for $SB_{\alpha\beta}$ at time slot t_{i+2} is $TP_{\alpha\beta} = max(TP(\alpha_{t_{i+2}}) + TP(\beta_{t_{i+2}}))$ where $D(\alpha_{t_{i+1}})$ is the traffic demand for SP_α at time slot t_{i+1} and $TP(\alpha_{t_{i+2}})$ is the predicted traffic demand for SP_α at time slot t_{i+2}. For example,

in Table 5.2, SPs between node-pairs (A, C) and (C, E) are banded with SP between node-pairs (A, E) since $B_f = 1$. For this band, the traffic demand at time slot t_{i+1} is $D_{\alpha\beta_{t+1}} = 5$ and the predicted traffic demand at time slot t_{i+1} is $TP_{\alpha\beta_{t+1}} = 5$. However, if these SPs are not banded together, then PPSA would assign 3 sub-carriers to each of the SPs between node-pairs (A, C), (C, E) and (A, E) requiring a total of 6 sub-carriers. By forming the spectrum band, only 5 sub-carriers are enough for these SPs. Once all the possible SPs are banded, we can use PPSA to assign sub-carriers to banded SPs and other SPs. The detailed explanation of PSBA is given in Algorithm 6.

Algorithm 6 Predictive Spectrum Band Allocation (PSBA)

Input: $\{SA(t_i, s, d)\}$, $D\{(t_{i+1}, s, d)\}$ and $\{TP(t_{i+2}, s, d)\}$
Output: $\{SA(t_{i+1}, s, d)\}$

1: Sort the traffic demands in the descending order of sub-carriers requested
2: **for** each SP α in $D(t_{i+1}, s, d)$ **do**
3: Calculate B_f of α with the neighboring SPs
4: Based on B_f, band SP_α with SP_β as spectrum band $SB_{\alpha\beta}$
5: Update the traffic demand of $SB_{\alpha\beta}$ as $max(D(\alpha_{t_{i+1}}) + D(\beta_{t_{i+1}}))$ and the predicted traffic demand as $max(TP(\alpha_{t_{i+2}}) + TP(\beta_{t_{i+2}}))$ at time slots t_{i+1} and t_{i+2} resp.
6: **end for**
7: Apply PPSA algorithm

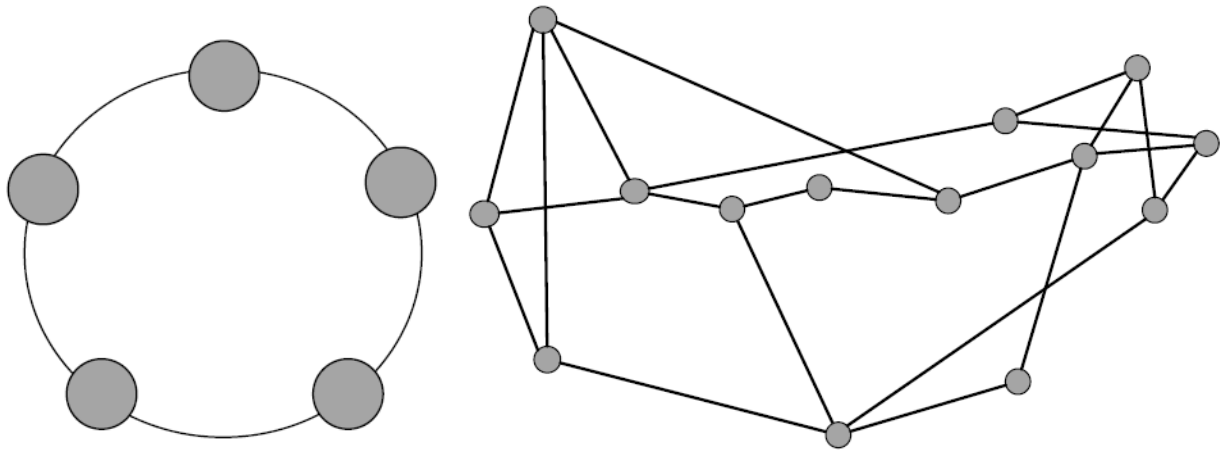
5.6 Performance Evaluation of Proposed Spectrum Allocation Schemes

We evaluate the performance of SAS, PSA, PPSA and PSBA in the 5-node ring network as shown in Figure 5.8(a) and in the 14-node NSFNET as shown in Figure 5.8(b). Since the ILP model proposed in Section 5.4 is computationally expensive for a large network, we evaluate the performance of the ILP model only in the small 5-node ring network. If not specified, the future traffic is predicted as in [80].

5.6.1 5-node Ring Network

A 5-node ring network is used to compare the performance of the ILP model with other heuristics. The ILP model is implemented using the CPLEX 12.2 [81]. The number of

connection is randomly generated within the range [5, 10]. For each connection, the source and the destination are picked at random and the sub-carrier demand for each connection is randomly generated within the range [1, λ] where λ denotes the traffic intensity. The sub-carrier demands for connection requests vary randomly in the successive time slots. In fact, in real networks, it can be observed that the traffic changes are correlated. For example, the traffic grows during day hours or decreases during night hours at similar time slots for different connections. However, due to the lack of access to such realistic traffic variations, we assume random traffic variations. We set the number of sub-carriers in each link high enough (i.e., 100) to ensure that every demand can be satisfied in this 5-node ring network (i.e. Eq. (5.3) in ILP is satisfied) and there is no blocking. The number of time slots in a period is set to 5. When running the ILP model, we set Eq. (5.1) as the primary objective to obtain the results in Table 5.3, and we use Eq. (5.2) as the primary objective to achieve the results in Figure 5.9. The simulation is performed a number of times and the average values are presented.



(a) 5-node ring network (b) 14-node NSF network

Figure (5.8) Networks used in the performance evaluation.

Table 5.3 shows the values of MS which is the maximum index of the sub-carriers allocated among all the fibers in a ring network for different values of λ. From the table,

Table (5.3) The values of the maximum index of the sub-carriers (MS) when applying Eq. (5.1) for different λ.

λ	ILP	SAS	PSA	PPSA	PSBA
3	24	25	28	27	27
4	35	36	39	38	37
5	43	45	49	48	48

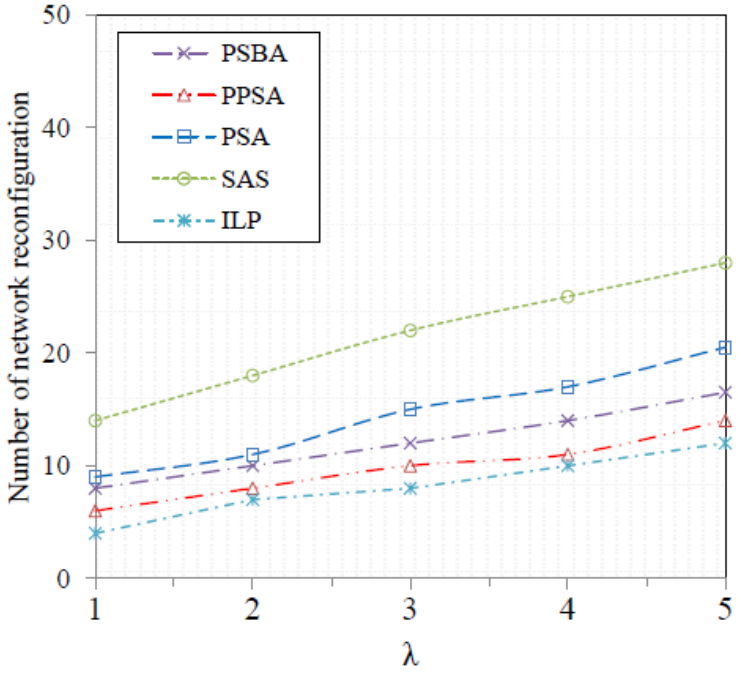


Figure (5.9) Number of network reconfiguration for the heuristic algorithms and the ILP model when applying Eq. (5.2).

we can see that the values of MS in PSA, PPSA and PSBA are larger than the values obtained from the ILP model. The values of MS in SAS are slightly greater than the values obtained from the ILP model. In terms of the MS, SAS outperforms other three heuristics. This is because in each time slot, SAS assigns sub-carriers to each SP from scratch without considering the disruption to existing connections and spectrum reservation for the future traffic. As a result, it can make efficient use of the spectrum and keep MS low in the network.

Figure 5.9 shows the performance of the ILP model and the heuristic algorithms in

terms of the number of network reconfiguration[2] when increasing the traffic intensity. From the figure, it can be seen that the number of network reconfiguration increases with the increase in the traffic intensity for all the schemes. Clearly, increasing the traffic intensity requires more traffic to be added or dropped resulting in the increase of the number of network reconfiguration in each scheme. The number of network reconfiguration in PSBA and PPSA are very close to the solution obtained from the ILP model while the number of network reconfiguration in PSA is slightly higher than that from ILP, PSBA and PPSA. SAS performs worse than the other three heuristics. This is because at any time slot, the naive SAS scheme assigns sub-carriers to SP from scratch without considering the existing spectrum assignment and the future traffic demands. Hence, SAS may frequently tear down existing connections and set-up new connections, resulting in high network reconfiguration cost.

Table (5.4) Running time for different algorithms in the 5-node ring network

Algorithms	Running time (mins)
ILP	20 - 25
SAS	0.5 - 1
Other heuristics	2 - 4

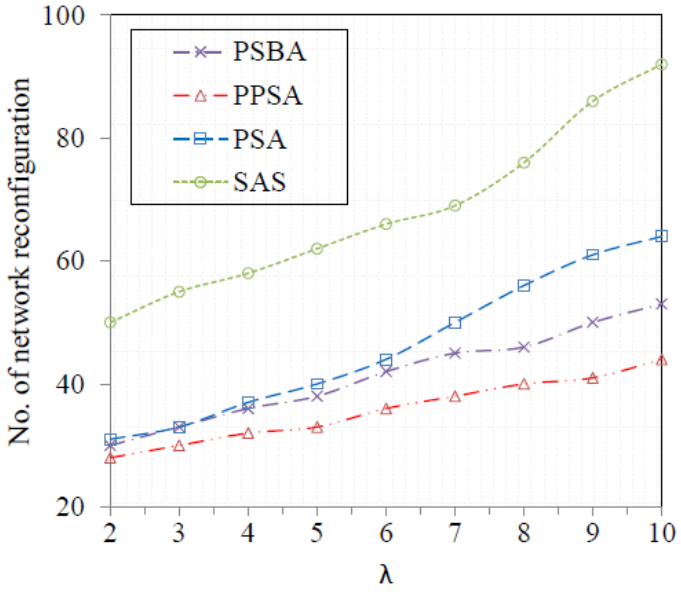
Table 5.4 shows the running time for different algorithms when simulating the 5-node ring network using a machine with 1.7 GHz Intel Core i3-3120M processor and 4 GB RAM. Since the proposed ILP model is intractable in a large network such as the 14-node NSFNET, we only compare the performance of the proposed heuristics.

5.6.2 14-node NSFNET

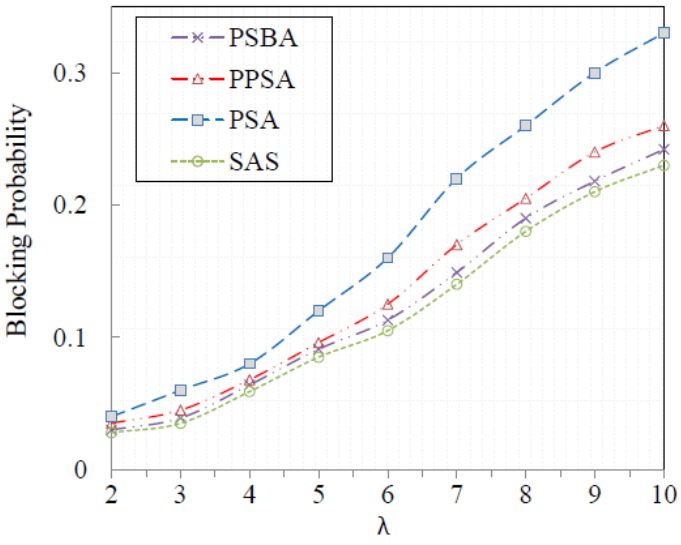
For the evaluation of the proposed schemes in the 14-node NSFNET, we assume that the time period (P) is divided into 12 time slots and there are 100 sub-carriers in each fiber. The number of connections is randomly generated within the range [15, 30] and the sub-carrier

[2]Note that, the simulation only considers the spectrum reconfiguration along the same path.

demand for each connection is randomly generated within the range [1, λ] where λ denotes the traffic intensity. The sub-carrier demands between connection requests vary randomly in the successive time slots.



(a) λ vs. Number of network reconfiguration



(b) λ vs. Blocking probability

Figure (5.10) Number of network reconfiguration and blocking probability in the 14-node NSFNET.

Figure 5.10(a) shows the average number of network reconfiguration when increasing λ. As the traffic intensity increases, the number of network reconfiguration increases in all four schemes. As the traffic intensity increases, PSA, PPSA and PSBA outperform SAS by a considerable margin. For example, when $\lambda = 5$, the number of network reconfiguration in SAS is 33% more than the other schemes. This is primarily because SAS assigns sub-carriers to the traffic demands without considering the existing spectrum assignments and the future traffic demands. Hence, in order to accommodate fluctuating traffic demands, SAS may frequently allocate a different set of sub-carriers to existing connections, which may require a significant amount of network reconfiguration. Unlike SAS, PSA, PPSA and PSBA consider the current spectrum assignments in addition to the traffic demands. Moreover, PPSA and PSBA consider the prediction of traffic demands at future time slots to assign sub-carriers which results in even lower network reconfiguration cost. Interestingly, PPSA outperforms PSBA indicating the potential ill-effect of banding/grouping of SPs in PSBA. In fact, when SPs with opposite change in spectrum demands are banded together in PSBA, this may lead to relocation of both SPs if any one of the banded SPs has to be relocated. This may cause a slight increase in network reconfiguration cost in PSBA when compared with PPSA.

In Figure 5.10(b), we limit the maximum number of available sub-carriers per fiber to 50 and compare the blocking probability when increasing λ. Here, blocking probability is the ratio of the number of connections that are blocked to the total number of connection requests. It can be seen from the figure that the blocking probability increases for all the four schemes with the increase of traffic intensity. From the figure, one can further see that SAS slightly outperforms PSBA and PPSA and PSA is the worst in terms of the blocking probability. This is because SAS assigns sub-carriers to SPs from scratch without considering the existing spectrum assignments. Hence, SAS makes better use of the spectrum at the cost of high network reconfiguration. To reduce the reconfiguration/disruption cost, PPSA may reserve some sub-carriers for future SPs and thus, resulting in a higher blocking probability than that of SAS. When comparing with PPSA, PSA causes a higher number of reallocation, which indicates that more connections are moved to a different set of sub-carriers. After this

moving/relocation, the previously assigned sub-carriers are now vacant and may or may not be sufficient to establish new connections. Hence, PSA yields more blocking than PPSA. PSBA also attempts to reserve the extra sub-carriers left vacant by the shrinking SPs for the possible expansion of the neighboring SPs. This increases the spectrum efficiency in PSBA leading to the decrease in blocking probability when compared to PPSA.

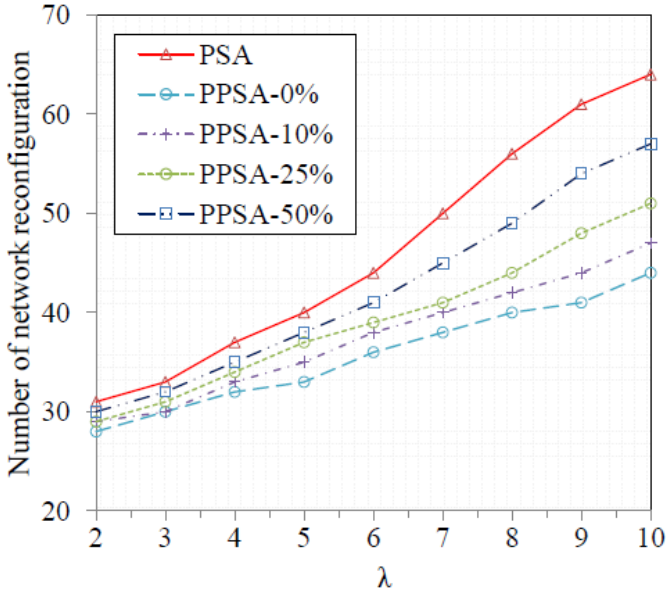


Figure (5.11) Number of network reconfiguration for different margin of traffic prediction errors.

As the performance of PPSA depends upon the predicted traffic demands, we study the impact of the accuracy of the predicted traffic demands on the performance of PPSA. For this, we simulate PPSA with different prediction accuracy in the 14-node NSFNET. We introduce noise by allowing the future traffic demands to fluctuate from the predicted traffic demands. In particular, the margin of errors from the predicted traffic demands considered in this simulation are 0%, +/-10%, +/-25% and +/-50% which are denoted by PPSA-0%, PPSA-10%, PPSA-25% and PPSA-50%, respectively. The number of reconfiguration for different margin of errors is shown in Figure5.11. As seen from the figure, the amount of network reconfiguration increases with the increase in the prediction error. The better

traffic prediction requires fewer network reconfiguration for PPSA. As the prediction error increases, the number of sub-carriers reserved by PPSA for future traffic change in successive time slots may not be enough, which results in the performance degradation of PPSA. In our simulation, similar pattern can be seen in PSBA as well since the performance of PSBA also depends on PPSA.

5.7 Chapter Summary

Since Internet traffic follows a similar pattern over a certain period of time, we can use this periodic behavior of Internet traffic to efficiently allocate resources in Spectrum-sLIced Elastic optical path networks (EONs) or Flex-Grid networks. In this chapter, we have proposed an Integer Linear Programming (ILP) model to formulate the time-varying traffic assignment problem in EONs or Flex-Grid networks. Since the ILP model is intractable for large networks, we have also proposed novel spectrum assignment schemes based on interference graph, namely Planning-ahead Spectrum Allocation (PSA), Predictive Planning-ahead Spectrum Allocation (PPSA) and Predictive Spectrum Band Allocation (PSBA). PSA considers the existing spectrum allocation and the traffic requests during the process of sub-carrier allocation to minimize the number of spectrum reconfiguration. PPSA adopts a plan-ahead mechanism whereby the spectrum allocation process not only considers the traffic demands in the current time slot, but also considers the existing spectrum assignment as well as the traffic prediction for the future time slot. Using this approach, PPSA can prevent unnecessary network reconfiguration at future time slots. In PSBA, spectrum paths (SPs) are banded/grouped together based on their band factor (B_f) before applying PPSA scheme for the spectrum allocation. Simulation results have shown that in comparison to PSA, PPSA and PSBA can achieve much less network reconfiguration while maintaining a low blocking probability for the time-varying traffic.

PART 6

APPLICATION OF ELASTIC OPTICAL NETWORKS IN NETWORK VIRTUALIZATION

Recent innovations in network virtualization and elastic optical networks (EONs) enable flexible deployment of optical networks as a service. However, one open challenge is how to embed virtual network (VN) requests onto the physical substrate network to maximize the sharing of physical resources, which is the so called virtual network embedding (VNE) problem. This chapter presents a novel virtual network embedding algorithm called alignment and consecutiveness-aware virtual network embedding (ACT-VNE), which takes into account the spectrum alignment and consecutiveness between adjacent fiber links when mapping virtual nodes/links onto the physical substrate nodes/links. This chapter also introduces a relative importance factor in ACT-VNE as importance, alignment and consecutiveness-aware virtual network embedding (iACT-VNE) that assigns virtual nodes to substrate nodes within close proximity to establish connection with less bandwidth utilization. Since EONs are prone to defragmentation, this chapter also proposes a min-max reconfiguration scheme called relative consecutiveness loss-aware and misalignment-aware virtual network reconfiguration (RCLM-VNR) that minimizes relative consecutiveness loss and maximizes alignment with adjacent links when reconfiguring the virtual network. Simulation results show that ACT-VNE and RCLM-VNR yield a lower blocking probability and a higher link utilization ratio, which leads to better utilization of the physical resources and increased revenue.

6.1 Network Virtualization

Recent advancements in the areas of mobile technologies, data center networks, cloud computing and social networks have triggered the growth of a wide range of network applications. The data rate of these applications vary from a few megabits per second (Mbps) to

Gigabits per second (Gbps) increasing the burden on the Internet. To efficiently handle such a huge amount of data traffic, and to cater to every business needs, network virtualization has been proposed as a promising solution [82]. With network virtualization, the functions of a traditional Internet service providers (ISPs) can be separated into two separate entities, infrastructure providers (InPs) and service providers [82]. InPs manage the physical infrastructures and resources while service providers can lease physical resources from one or more InPs according to the need and provision a virtual network (VN). Network virtualization allows service providers to dynamically compose multiple heterogeneous virtual networks that co-exist together in isolation form. This enables the use of Network as a service (NaaS) as per the need and also, help mitigate the Internet ossification problem. A typical VN is composed of several virtual nodes interconnected by virtual links. Each virtual node requires a certain amount of computing resources from the substrate node while each virtual link is a path consisting of one or multiple substrate links with a certain amount of bandwidth. The process of mapping virtual nodes/links to substrate nodes/links is called virtual network embedding (VNE) problem, which is NP-hard [83].

Recently, Flex-Grid or Elastic optical networks (EONs) is viewed as one of the promising substrate technologies. EONs employ the orthogonal frequency division multiplexing (OFDM) technology to distribute data on just enough sub-carriers. The adjacent sub-carriers in EONs can overlap each other and EONs can support different bit rates using various modulation schemes [7], [6]. The use of EONs bring additional flexibility in terms of more granular sub-carriers and more flexible bandwidth assignment [1], [6]. At the same time, the use of EONs bring additional constraints such as *sub-carrier consecutiveness* and *spectrum-continuity* constraints while assigning a spectrum path (SP) to any virtual link demand. Based on whether the connections in VN can be set-up all optically in the substrate network or not, VNE in EONs can be classified into opague VNE (o-VNE) and transparent VNE (t-VNE) [84], [85]. In t-VNE, all the virtual links in VNs must use the same spectrum on substrate fiber links to ensure all the connections in a VN can be established all-optically (in the absence of spectrum converters) while in o-VNE, different virtual links can employ

different sets of sub-carriers.

6.1.1 Related works of VNE in EONs

The authors in [84] formulated VNE in the context of establishing connections all optically and provided an Integer linear programming (ILP) formulation for the link mapping in both *o*-VNE and *t*-VNE in WDM networks. However, the ILP model presented in this paper did not consider the node mapping. In [86], the authors considered the link mapping in EONs but did not address the issue of node mapping. The authors in [87] proposed an ILP method and two heuristics algorithms for *o*-VNE in EONs where multiple virtual nodes can be mapped to the same substrate node. The network condition changes over time due to the random arrival/departure of VNs and the resources at the substrate network may get fragmented. Hence, to capture this change in resource capacity, the work in [72] proposed a model for *o*-VNE using alignment and consecutiveness aware node rank. The authors in [88] introduced spectrum consecutiveness to measure the capacity of substrate fiber links. The authors in [85] used spectrum consecutiveness and the capacity of the substrate node to form the local resource capacity (LRC) of substrate nodes. The authors also proposed an ILP model and heuristic algorithms using LRC of a substrate node for both *o*-VNE and *t*-VNE. While the spectrum consecutiveness can measure the available capacity of a spectrum band or a single fiber link, it may not present a true capacity of a substrate node.

Particularly, for *t*-VNE in EONs where all the virtual links of a VN should be assigned with the same set of sub-carriers, the LRC used in [85] may not work in *t*-VNE efficiently. For example, Figure 6.3 shows one example of a substrate network consisting of fours nodes *A, B, C* and *D* with a computational capacity of 20, 20, 10 and 10, respectively. There are 10 sub-carriers in each substrate fiber link (i.e., *AB*, *BC*, *CD*, *AD*). When the LRC scheme from [8] maps the VN request in Figure 6.1(b), virtual nodes a, b and c is mapped to substrate nodes A, B and C resp. In this case, the link mapping would fail as there are not enough free sub-carriers of the same index as required in *t*-VNE. However, if the node mapping is done as given in Figure 6.3(d), the links AD and CD can be assigned sub-carriers with the same

index. Accordingly, in this work, we introduce the alignment and consecutiveness-aware node ranking to measure the node capacity more accurately. In addition, we also propose a novel relative importance factor in t-VNE that attempts to map virtual nodes to substrate nodes with close proximity (to avoid the potential long routing paths).

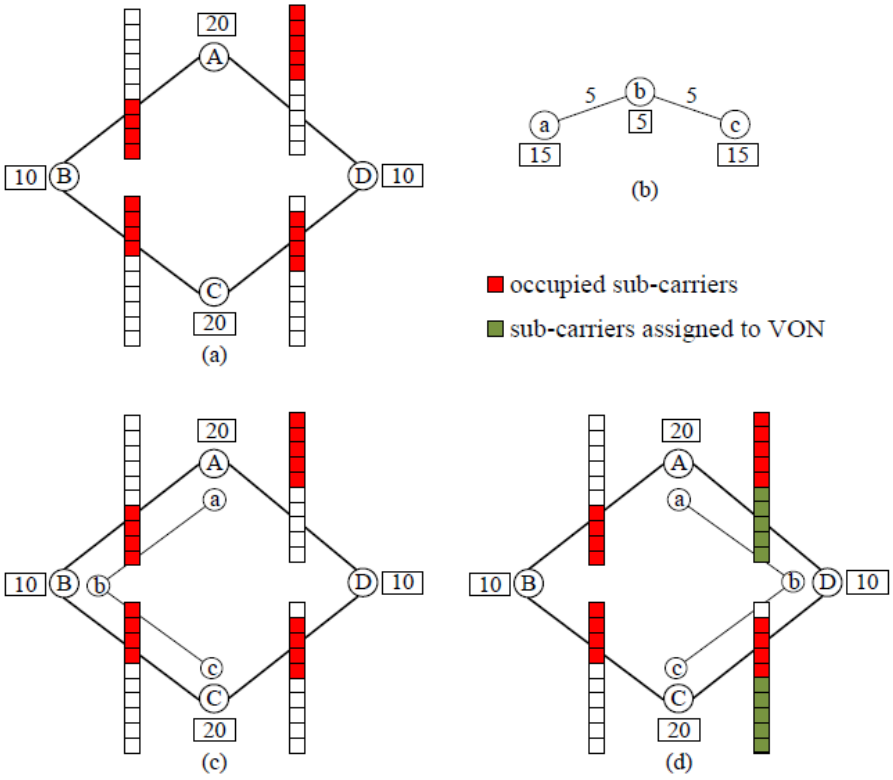


Figure (6.1) (a) Subtrate Network (b) VN requests (c) Failed VNE (d) Successful VNE

6.2 Network Model and Problem Formulation

A VN request can be modeled as an undirected weighted graph $G_V = \{N_V, L_V\}$, where N_V represents the set of virtual nodes and L_V represents the set of virtual links. We use $c_r(v)$ as the computing requirement of node v and $b_r(uv)$ as the bandwidth requirement of virtual link uv. Similarly, the substrate network can be modeled as an undirected weighted graph $G_S = \{N_S, L_S\}$, where N_S represents the set of substrate nodes and L_S represents the set of substrate fiber links. The available computing capacity of a substrate node V is

$C_c(V)$ and the available bandwidth capacity of a substrate fiber link UV is $B_c(UV)$. The transparent virtual network embedding (t-VNE) problem in EONs can be defined as follows.

t-VNE Problem - Given a VN request $G_V = \{N_V, L_V\}$ and substrate network $G_S = \{N_S, L_S\}$, can the VN be mapped onto the substrate network while satisfying the following requirements: (i) a virtual node v must be mapped to only one substrate node V such that $c_r(v) \leq C_c(V)$; (ii) when virtual nodes u and v are mapped to substrate nodes U and V, respectively, the virtual link uv is mapped to a spectrum path between U and V following the constraints of *sub-carrier consecutiveness* and *spectrum-continuity* and $b_r(uv) \leq B_c(UV)$; (iii) all the virtual links of the VN should be assigned with the same set of sub-carriers?

The VNE problem in WDM networks is proven to be NP-hard [89]. The VNE in EONs consists of an instance of VNE in WDM networks, which makes t-VNE NP-hard. The requirement of assigning all the virtual links with the same set of sub-carriers makes t-VNE even more complicated. Hence, in this chapter, we propose two heuristic algorithms for t-VNE, namely Alignment and consecutiveness-aware transparent virtual network embedding (ACT-VNE) and important, alignment and consecutiveness-aware virtual network embedding (iACT-VNE). In the next section, we explain the alignment and consecutiveness-aware (ACT) node ranking scheme that will be used by both ACT-VNE and iACT-VNE.

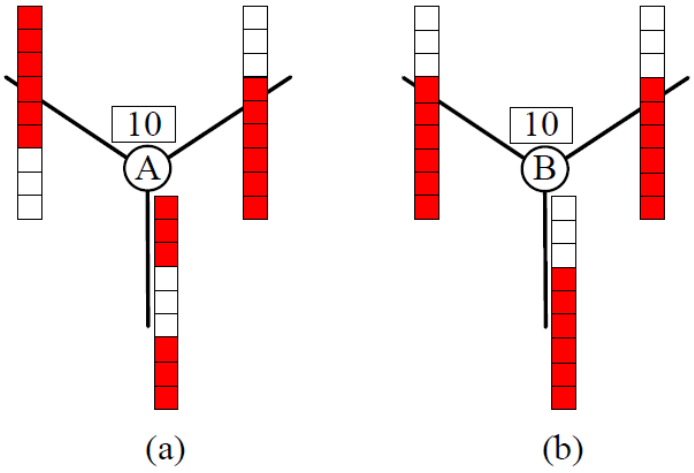


Figure (6.2) (a) Node A (b) Node B with different empty sub-carriers on adjoining links

6.3 Alignment and ConsecuTiveness-aware Node Ranking

For a network $G(V, E)$ where the computing capacity of a node v is $C_c(v)$ and the bandwidth capacity of a link e is $B_c(e)$, the Local Resource Capacity (LRC) [83] of node v is defined as follows.

$$h_v = C_c(v) \sum_{(v,u)\in E} B_c(vu), \quad \forall v \in V \tag{6.1}$$

Eq. (6.1) can be modified to measure the consecutiveness-aware local resource capacity (CaLRC) [85] of a substrate node v_s in substrate network $G_S = \{N_S, L_S\}$, as follows:

$$h_{v_s} = C_c(v_s) \sum_{(v_s,u_s)\in E_S} q(v_s, u_s), \quad \forall v_s \in N_S \tag{6.2}$$

where $q(v_s, u_s)$ is the spectrum consecutiveness of the adjoining substrate fiber link. The spectrum consecutiveness of the substrate fiber link, $e_s = (v_s, u_s)$ [88] can be calculated as follows:

$$q_{e_s} = \sum_{k} \sum_{i=1}^{n_{e_s}^k} (n_{e_s}^k - i + 1) \cdot \alpha_i \quad \forall e_s = (v_s, u_s) \in L_S \tag{6.3}$$

where k is the number of spectrum blocks in substrate fiber link e_s, $n_{e_s}^k$ is the number of available sub-carriers in the k^{th} spectrum block in e_s and α_i is the weight co-efficient. The value of α_i is equal to 1 if the spectrum demand includes block of size i and otherwise 0.

While CaLRC may be used for node mapping in o-VNE, it may not be applicable in t-VNE for a couple of reasons. First, any node V with a high computational capacity may have a high rank even if the value of spectrum consecutiveness is as small as one in all of V's adjoining links. Any virtual node with virtual links that require more than one sub-carrier in substrate fiber links (SFLs) may be blocked if mapped to this node. A better node ranking should be relative to the incoming VN request and adaptive as the resources in the substrate network change after accommodating each VN request. The second issue is that more consecutive sub-carriers on substrate fiber links may not translate to more available spectrum paths in EONs without spectrum conversion. This is because a spectrum

path corresponding to the virtual link often travels multiple hops in the substrate network. For example, Figure 6.2(a) and (b) show the substrate node A and B having 10 units of computational capacity and three adjoining links with 3 free sub-carriers in each link. Ranking node A and B on the basis of CaLRC will give them equal ranks. However, no traffic can pass through node A without undergoing spectrum conversion and hence, any virtual node mapped to substrate node A will be blocked as the adjoining links of node A are not spectrally aligned with each other. However, Figure 6.2(b) can satisfy a demand with 3 or less sub-carriers going through node B. Hence, it is important to consider the alignment of the available consecutive sub-carriers along adjacent substrate fiber links. In this section, we propose a sub-carrier alignment factor that measures the sub-carrier alignment among adjoining links and calculate the node rank relative to each incoming VN request. This will give the lower bound on the number of sub-carrier that this node can handle without spectrum conversion. Specifically, for a node v between two adjacent links uv and vu', the sub-carrier alignment factor of node v is defined as follows:

$$a_v = N - \sum_{i=1}^{N} e_{s_i}^{uv} \vee e_{s_i}^{vu'} \tag{6.4}$$

where N is the number of sub-carriers in the links; $e_{s_i}^{uv} = 1$ if the sub-carrier s_i is occupied in link uv, otherwise $e_{s_i}^{uv} = 0$; and $\vee$ is the logical OR operation. The a_v for node A in Figure 6.2(a) is 0 meaning that no more traffic can pass through node A. However, a_v for node B in Figure 6.2(b) is 3 meaning that node B can allow any demand with 3 or less sub-carriers to pass through. Moreover, a node with high spectrum alignment factor can assign the same set of sub-carriers to a VN along the incoming and the outgoing links, eliminating any need for spectrum conversion at the node.

Here, we propose an alignment and consecutiveness-aware node ranking for t-VNE, namely ACT node rank, which uses the capacity of the substrate node along with the spectrum consecutiveness and spectrum alignment of the adjoining links relative to the incoming

VN request to calculate the node rank. The ACT node rank is given as follows:

$$ACT\text{-}Rank(v_s) = C_c(v_s) \cdot a_{v_s} \cdot q(v_s), \quad \forall v_s \in V_S \tag{6.5}$$

where $C_c(v_s)$ is the computing capacity, a_{v_s} is the spectrum alignment factor and $q(v_s)$ measures the spectrum consecutiveness in all the adjoining links of substrate node $v_s \in V_S$. For an incoming VN request G_V, $q(v_s)$ is given as follows:

$$q(v_s) = \sum_{e_s=(v_s,u_s)} \sum_{b_{e_s}} \sum_{i=min\{L_V\}}^{n^s_{b_e}} (n^s_{b_e} - i + 1) \cdot \alpha_{b_e}, \; {}^{\forall v_s \in N_S,}_{\forall e_s \in L_S} \tag{6.6}$$

where $min\{L_V\}$ is the (projected) minimum bandwidth requirement in terms of number of sub-carriers for the incoming VN request, $n^s_{b_e}$ is the size of the spectrum block b in SFL e and α_{b_e} is the relative weight coefficient. The value of α_{b_e} indicates the weight of each spectrum block b in SFL e relative to the incoming VN request G_V, which is calculated as below:

$$\alpha_{b_e} = \frac{n^s_{b_e}}{max\{L_V\}} \tag{6.7}$$

where $max\{L_V\}$ is the (projected) maximum number of sub-carrier required in all the links of the incoming VN request G_V. If the spectrum block b in SFL e does not have $max\{L_V\}$ consecutive sub-carriers, then $\alpha_{b_e} < 1$ decreasing the rank of the corresponding substrate node. The relative weight coefficient makes ACT node rank adaptive to the requirement of the incoming VN request. Similarly, the sub-carrier alignment factor of the node v_s in Eq. (6.5) can be calculated as follows:

$$a_{v_s} = \sum_{\substack{(v_s,u_s)\in L_S,\\ u_s \neq u'_s}} \frac{(N - \sum_{i=1}^{n}(e^{u_s v_s}_{s_i} + e^{v_s u_{s'}}_{s_i})}{N} \tag{6.8}$$

where N is the number of sub-carriers in SFL (v_s, u_s), $e^{u_s v_s}_{s_i} = 1$ if the sub-carrier slot s_i is occupied in the SFL (v_s, u_s) and 0 otherwise.

6.4 Alignment and Consecutiveness-aware Transparent Virtual Network Embedding (ACT-VNE)

In this section, we present our alignment and consecutiveness-aware transparent virtual network embedding (ACT-VNE) algorithm. ACT-VNE uses ACT node ranking for node mapping. For link mapping, it uses K-shortest-path routing and first-fit spectrum assignment.

The details of ACT-VNE algorithm are explained in Algorithm 7. In Line 1, for each incoming VN request, the LRC of the virtual node is measured as given in Eq. (6.2) and then the virtual nodes are ranked according to the LRC. In Line 2, the ACT node rank is calculated for each substrate node N_S in G_S relative to G_V using Eq. (6.5). In Line 3, for each VN request, substrate nodes are ranked according to the ACT node rank. The node mapping is done in Lines 5 - 11. Whenever the virtual node's computing resource requirement can be satisfied, the virtual node with the largest LRC is mapped onto the unmapped substrate node with the largest ACT node rank. If not, the virtual network is blocked. Then, the link mapping is done in Lines 12 - 24. For the link mapping, the virtual links are first sorted based on their bandwidth requirements in descending order, and then spectrum allocation is performed for each virtual link with the K-shortest-path routing and first-fit (FF) spectrum assignment.

6.5 Importance, Alignment and Consecutiveness-aware Transparent Virtual Network Embedding (iACT-VNE)

ACT-VNE uses ACT node ranking that defines the global rank of a substrate node in a substrate network. If VNE is done solely based on ACT node rank, two adjacent virtual nodes may be mapped to two substrate nodes that are far apart from each other. This may require a longer substrate spectrum path and increase the usage of bandwidth resources between two substrate nodes.

Hence, in this section, we introduce the concept of the relative importance of one node

Algorithm 7 Alignment and Consecutiveness-aware Transparent Virtual Network Embedding (ACT-VNE)

Input: Substrate Network G_S, a VN request G_V
Output: Node mapping N_M, Link Mapping L_M

Rank N_V in G_V according to LRC in Eq. (6.2)
Calculate *ACT-Rank* for N_S in G_S relative to G_V as given in Eq. (6.5)
Rank N_S according to the ACT node rank
$Q_V \leftarrow$ sorted virtual nodes in G_V according to LRC
for $i = 0$ to $|Q_V| - 1$ **do**
    Map $Q_V[i]$ to unmapped substrate node in N_S with the highest ACT node rank.
    if node-mapping fails **then**
        Mark G_V as blocked
        break;
    end if
end for
if G_V is not marked as blocked **then**
    Sort L_V on the basis of bandwidth demand
    for all $e_v \in L_V$ **do**
        Find k-shortest paths for e_v in G_S
        for each path in k-shortest paths **do**
            if there are enough empty sub-carriers **then**
                Use first-fit scheme for spectrum assignment
            end if
        end for
        if spectrum assignment fails **then**
            Mark G_V as blocked
            break;
        end if
    end for
end if

to other nodes using the number of paths that connect them. The rank of nodes in the graph should vary after assigning a virtual node to a substrate node. The nodes that are in near proximity to the assigned node should have higher rank than the nodes that are far away. To capture this, we introduce the relative node importance in VNE that will encourage the mapping of virtual nodes to substrate nodes in close proximity. We use the definition of node importance as given in [90] and combine this with ACT-VNE to form the proposed Importance, Alignment and Consecutiveness-aware Transparent Virtual Network Embedding (iACT-VNE). The importance of node v with respect to node n defined as $I(v/n)$ [90] can be calculated as follows:

$$I(v/n) = \sum_{i=1}^{|P(v,n)|} \beta^{-|p_i|} \tag{6.9}$$

where $P(v, n)$ is a set of K-shortest node disjoint paths from v to n, p_i is the i^{th} path in P, λ is a scalar co-efficient, $1 \leq \beta \leq \infty$ determines how much importance is conferred from v to n. Here, we take the value of $\beta = 2$ as given in [90]. Here, two nodes are related according to the path that connect them and the longer a path is, the less importance is conferred along that path. We consider K-short node-disjoint paths that have neither edges nor nodes in common. In other words, no node or edge can be used more than once in the set of multiple paths from v to n.

Given a VN, $G_V = \{N_V, L_V\}$, a set of remaining virtual nodes N_V^R in G_V and a set of assigned nodes in substrate network, N_S^A, the importance of a substrate node $v_S \in N_S \setminus N_S^A$ is calculated relative to all the nodes in N_S^A as shown in Eq. (6.10):

$$I(v_s/N_S^A) = \frac{1}{|N_S^A|} \sum_{n_s \in N_S^A} I(v_s/n_s) \tag{6.10}$$

Once, this importance v_s is known then it is multiplied with ACT rank to form the iACT node rank as shown in Eq. (6.11):

$$iACT\text{-}rank(v_s) = ACT\text{-}rank(v_s) \times I(v_s/N_S^A), \quad \forall v_s \in V_S \tag{6.11}$$

The steps for the proposed Importance, Alignment and Consecutiveness-aware Transparent virtual Network Embedding (iACT-VNE) is similar to Algorithm 7 except for Lines 2 and 3, where in Line 2, the iACT node ranking is calculated for each substrate node using Eq. (6.11) and in Line 3, the substrate nodes are ordered according to the calculated iACT node ranking.

6.6 Virtual Network Reconfiguration (VNR)

Given the dynamics in Internet, as the VN requests may come and leave randomly, the once efficient VNE mapping may not offer optimal resource utilization due to the fragmentation of spectral resources [91]. Hence, the incoming VN requests may have to be blocked even though there are sufficient sub-carrier resources available in the SFLs. Figure 6.3 shows one example of a substrate network consisting of three nodes *A, B, C*, each with 10 units of computational capacity and 10 sub-carriers at each SFLs (i.e., *AB*, *BC*, *AC*). Figure 6.3(b) shows three VN requests $(i), (ii)$ and (iii) that arrive in the order from top to bottom. Figure 6.3(c) shows the mapping for VN (i) and (ii) given by VNE following the *sub-carrier consecutiveness* and *spectrum-continuity* constraints. Now, if VN request (iii) has to be mapped, it will be blocked as there are not enough sub-carriers in SFLs *AB* and *AC*. However, if we reconfigure the VONs that are already mapped and move the sub-carriers assigned to virtual links *ac* on SFL *AC* to the higher spectrum, then all of the VONs can be mapped as shown in Figure 6.3(d). Hence, the blocking probability could be reduced by reconfiguring some of the already established virtual links/nodes. This reconfiguration of virtual links and virtual nodes is known as Virtual Network Reconfiguration (VNR).

6.6.1 Related works in VNR

While VNE is a very well-studied problem in WDM networks, there are only few papers that studies VNR schemes to optimize the resource utilization. In [92], the authors considered network virtualization for the first time in WDM networks and proposed a scheme that periodically checked for critical substrate fiber links or nodes based on Global marking

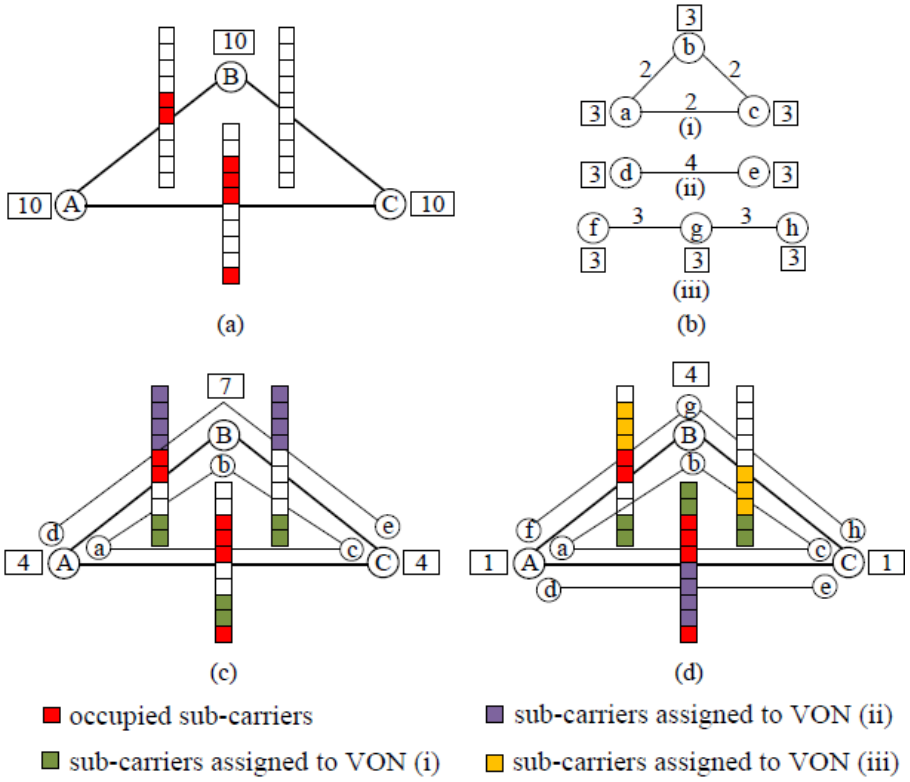


Figure (6.3) (a) Subtrate Network (b) VN requests (c) VNE results (a) VNE results after VNR

algorithm. If any such critical links or nodes are found, all virtual networks associated to these links or nodes are reallocated to a new physical nodes and links. The authors in [83] proposed schemes that periodically checks and recomputes virtual node links for active VN requests with longer residual lifetimes and allows substrate link network to split a virtual link over multiple substrate paths an employs path migration over these multiple paths to re-optimize the utilization of the substrate links. However, only virtual links can be reallocated which seriously limits its practicability. The authors in [93] proposed using the Critical Index and Popularity Index to determine the bottlenecks that causes the virtual node requests to be rejected and then moves these congested links or nodes to a less critical region. These schemes that periodically checks for resource scarcity are classified as proactive schemes and entails relatively high computational overheads and migration costs. Hence, reactive schemes are also proposed in literature that invokes VNR schemes after VN setup request fails. For

examples, the authors in [94] presents a VNR scheme that creates star moving candidate using a single virtual node and its associated virtual node links and migrates the one with the least migrations costs as opposed to migrating a whole virtual node topology. The authors in [95] proposed reactive reconfiguration scheme which performed multiple reconfiguration schemes in batch to minimize substrate link congestion while the authors in [96], the authors proposed a solution where various number of reconfiguration solutions are generated sequentially using genetic algorithm and selects the best reconfiguration solution that minimize both the migration and mapping cost.

In the literature, some work has been done for VNE and VNR in WDM networks [92] [93], [94] and [95]. However, these schemes cannot be applied directly for VNE and VNR in EONs or Flex-Grid networks because of the additional constraints such as the sub-carrier consecutiveness constraint [6]. In this chapter, we, for the first time, investigate VNR in EONs. We propose a novel VNE algorithm that does node ranking based on spectrum alignment and spectrum consecutiveness, and link mapping that minimizes relative loss in spectrum consecutiveness. In addition, we also present an efficient VNR algorithm in EONs that can be invoked when the incoming VN requests are blocked due to the lack of sub-carriers in SFLs.

6.7 Relative Consecutiveness Loss-aware and Misalignment-aware Virtual Network Reconfiguration (RCLM-VNR)

In *o*-VNE, the proposed ACT-VNE algorithm can be employed whenever there is a new incoming VN request. Since the VN requests arrive/leave dynamically, the spectrum resources may get fragmented [9]. Due to this fragmentation, the incoming VN requests may be blocked even though there are sufficient physical resources available in the substrate nodes. Denying services to VN requests even though there are sufficient spectrum resources is a huge loss to the service providers. Hence, in this section, we consider virtual network reconfiguration (VNR) in EONs. The objective of the VNR is to reduce the rejection rate of the incoming VN requests through on-demand virtual network reconfiguration. Since the

migration of virtual nodes may incur greater service disruption in comparison to the virtual link migration, and the blocking of the VN request is mostly due to the congested links (rather than exhausted nodes) [94], we only reconfigure the link mapping while keeping the node mapping intact. In specific, we propose a reactive min-max VNR algorithm called relative consecutiveness loss-aware and misalignment-aware virtual network reconfiguration (RCLM-VNR) [97].

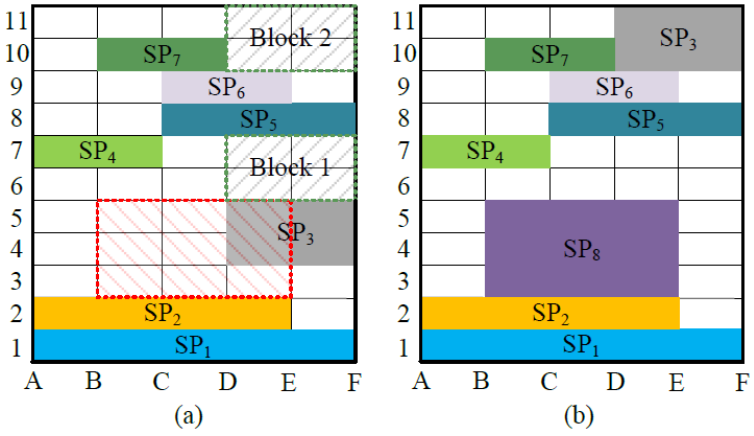


Figure (6.4) (a) Current state of the network (b) State after RCLM-VNR

RCLM-VNR reconfigures the spectrum assignment in SFLs when an incoming VN request cannot be satisfied due to the lack of sufficient consecutive sub-carriers in SFLs. The detailed steps for RCLM-VNR are explained in Algorithm 8. RCLM-VNR starts by finding K-candidate First-fit spectrum blocks for the blocked virtual link along the shortest path. Among the candidate spectrum blocks, RCLM-VNR will choose the block with minimum RCL when the established spectrum paths in the block are reallocated. The First-fit spectrum assignment scheme is then used to reallocate the already established spectrum paths. If two or more candidate blocks have the same RCL, then the one with the maximum spectrum alignment factor is chosen. If SP_b represents the set of spectrum paths that needs to be reallocated in candidate block b, then total RCL for block b can be calculated as follows:

$$RCL_b = \sum_{sp \in SP_b} RCL_p, \quad \forall b \in B \tag{6.12}$$

Algorithm 8 Relative Consecutiveness Loss-aware and Misalignment-aware Virtual Network Reconfiguration (RCLM-VNR)

Identify K candidate blocks B_K to establish spectrum path for the blocked virtual link along the shortest path
Sort the candidate blocks according to the minimum number of established spectrum paths
Initialize the variable for Discarded Blocks, $D_B \leftarrow 0$
for each candidate blocks **do**
  $RCL_{b_i} \leftarrow 0$
  for each spectrum path in the block **do**
    Generate k-candidate slot assignment sets, C_s
    if $|C_s| == 0$ **then**
      Discard this block
      $D_B \leftarrow D_B + 1$;
    else
      $RCL_{b_i} += min\{RCL(C_s)\}$
    end if
  end for
end for
if $K == |D_B|$ **then**
  Block the corresponding VN.
else
  Choose the block with minimum RCL_{b_i}
  if there is a tie **then**
    Choose the one that yields maximum alignment factor.
  end if
end if

An example of the VNR process is shown in Figure 6.4(a) where a virtual link, say *be*, requires a spectrum path, SP_8 consisting of 3 sub-carriers from node B to node E. This virtual link request and the corresponding VN request have to be blocked since there are not enough free consecutive sub-carriers along the path from B to E. However, when the proposed RCLM-VNR is employed, it will identify the potential blocks for this virtual link request as indicated by the red dash block in the figure. In order to make this red block available for the virtual link *be* request, spectrum path SP_3 has to be relocated. For the relocation, spectrum path SP_3 has two candidate locations as indicated by the green dash block 1 and 2. RCLM-VNR will ensure that spectrum path SP_3 is relocated to block 2 to minimize the total RCL loss in links DE and EF. Certainly, RCLM-VNR can be used to

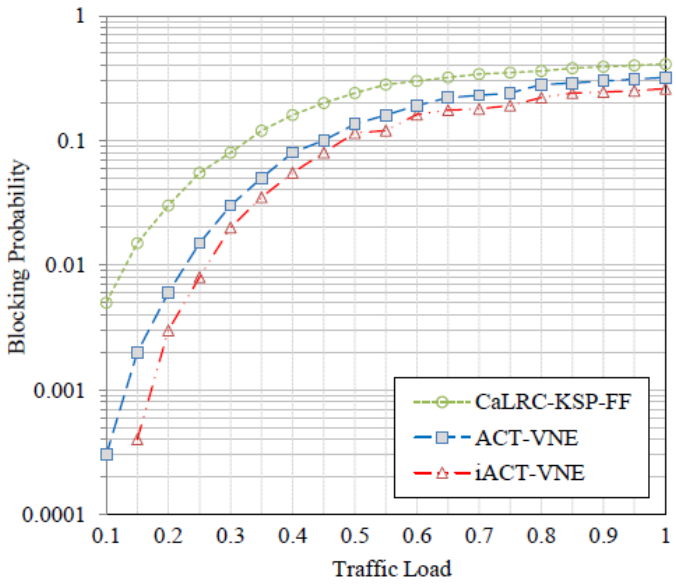


Figure (6.5) Blocking Probability vs. VN Request Intensity in a 20-node ARPANET

together with the proposed ACT-VNE when one of the virtual links is blocked due to the insufficient sub-carriers in the SFLs.

Table (6.1) Parameters used in the Simulation of VNE and VNR in EONs

Parameters	20-node ARPANET
Number of substrate nodes	20
Number of substrate fiber links	32
Computing capacity in a substrate node	500 units
Bandwidth capacity in a substrate link	200 units
Computing demand in a virtual node	[1,6] units
Bandwidth demand in a virtual link	[1,10] units
Number of virtual nodes in VN	[3,7]

6.8 Performance Evaluation of the proposed VNE and VNR algorithms

In this section, we compare the performance of the proposed ACT-VNE and iACT-VNE with Consecutiveness-aware Local Resource Capacity K-Shortest Path First-Fit VNE (denoted by CaLRC-KSP-FF from [85]). The scheme CaLRC-KSP-FF uses CaLRC from Eq.

(6.2) for node mapping and first-fit spectrum assignment on K-shortest path for link mapping. Table 6.1 shows the simulation parameters that are adopted in a 20-node ARPANET. The value of K is 3 for both K-shortest path algorithm and K-shortest disjoint paths. VN requests are generated according to a Poisson process with an average rate of λ, and each VN has an exponentially distributed lifetime. Each VN request has a random number of nodes and the virtual links between nodes are randomly created with a probability 0.5. The computing capacity of virtual nodes and bandwidth requirements of virtual links are randomly generated and VN requests are served in a first-come first-served basis.

Figure 6.5 shows the performance of ACT-VNE, iACT-VNE and CaLRC-KSP-FF in terms of blocking probability for a 20-node ARPANET. VN request intensity is defined as the ratio of the number of VN requests to the maximum number of VN requests used in the simulation, which is 50. It can be seen from the figure that the blocking probability increases with the increase in the VN request intensity for all the cases. In Figure 6.5, the blocking probability of ACT-VNE is marginally better than that of CaLRC-KSP-FF particularly for lower VN request intensity. This is because ACT-VNE uses the proposed ACT node ranking scheme, which considers spectrum alignment along with spectrum consecutiveness to identify proper node mapping for the incoming virtual node request. On the other hand, CaLRC-KSP-FF only takes the spectrum consecutiveness into consideration. iACT-VNE, which combines the factor of relative node importance along with ACT node ranking outperforms CaLRC-KSP-FF by a considerable margin and is marginally better than ACT-VNE. This is due to the fact that iACT-VNE additionally considers the importance of each node relative to the nodes already mapped, resulting in shorter substrate path length and less sub-carriers resource consumed. The performance gap between iACT-VNE, ACT-VNE and CaLRC-KSP-FF grows narrower for higher VN request intensity.

We also evaluated the performance of ACT-VNE, iACT-VNE and CaLRC-KSP-FF in terms of the average link utilization ratio as shown in Figure 6.6. The link utilization ratio is defined as the ratio of average number of sub-carriers used to the total number of sub-carriers in the link. From the figure, it can be seen that the average link utilization ratio

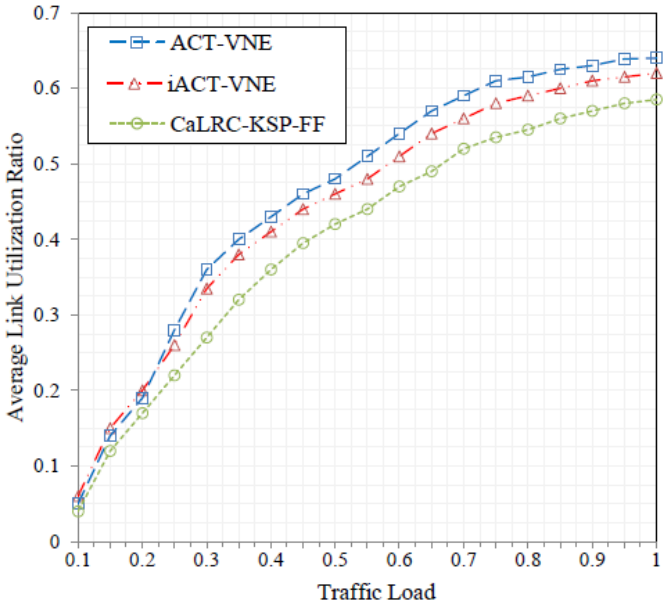


Figure (6.6) Average Link Utilization Ratio vs. VN Request Intensity in a 20-node ARPANET

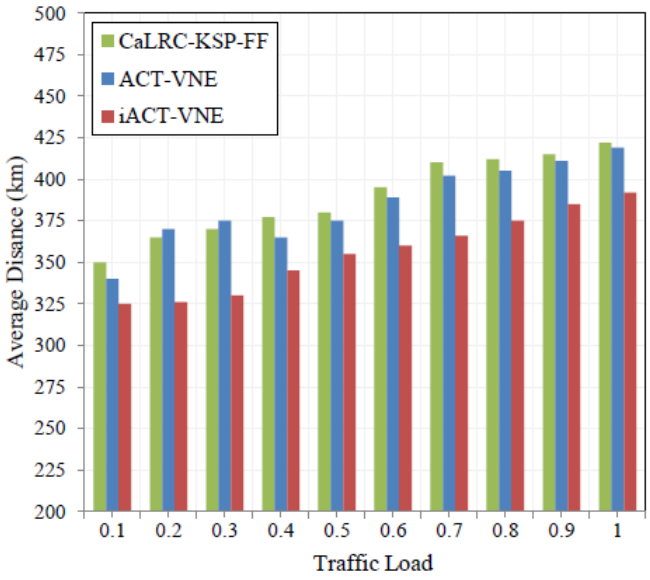


Figure (6.7) Average distance per embedded substrate path in a 20-node ARPANET

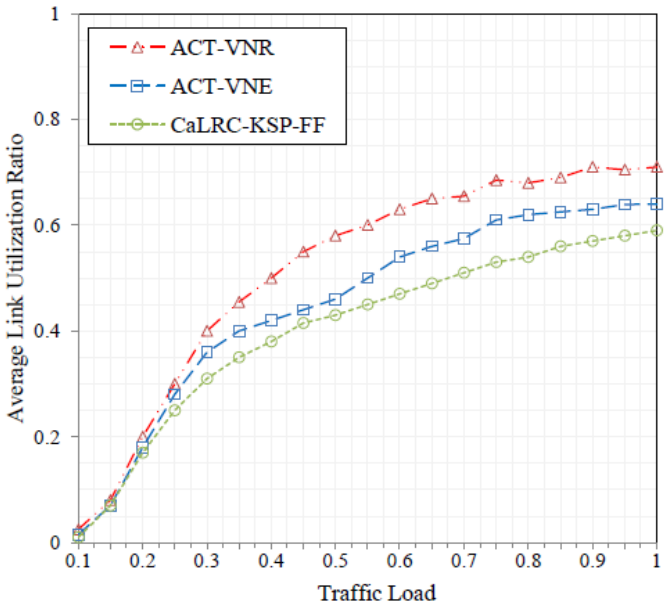


Figure (6.8) Average Link Utilization Ratio vs. Traffic Load in a 14-node NSFNET

increases with the increase in the VN request intensity for all the cases. The average link utilization in the proposed ACT-VNE and iACT-VNE is larger than that in CaLRC-KSP-FF since the proposed ACT-VNE can identify better node mapping and also, have less blocking probability than CaLRC-KSP-FF. The link utilization in iACT-VNE is very close to ACT-VNE as iACT-VNE assigns virtual nodes to substrate nodes that are in close proximity. The potential shorter routing paths and less sub-carrier resource consumption are preferred in iACT-VNE resulting in lower link utilization ratio in iACT-VNE. Note that iACT-VNE has lower blocking probability compared to ACT-VNE which indicates that iACT-VNE can accommodate more VN requests with less link utilization than ACT-VNE. The average path distance of each substrate path mapped for a virtual link is shown in Figure 6.7. The average path distance in iACT-VNE is consistently shorter than both CaLRC-KSP-FF and iACT-VNE. The reason is that iACT-VNE incorporates the proposed importance factor that prefers mapping virtual nodes to substrate nodes that are in close proximity. This result in shorter substrate path length and less sub-carrier resource consumption that can be clearly seen from the graph.

The performance of ACT-VNR, ACT-VNE and CaLRC-KSP-FF are also evaluated in

terms of the average link utilization ratio in a 14-node NSFNET as shown in Figure 6.8 using same parameters as in Table 6.4. From the figure, the average link utilization ratio increases with the increase in the traffic load in all cases. The average link utilization is larger in ACT-VNE than CaLRC-KSP-FF since the proposed ACT-VNE can identify better node/link mapping with less blocking probability. ACT-VNR outperforms both ACT-VNE and CaLRC-KSP-FF while taking advantages of the efficient node/link mapping process in ACT-VNE as well as the process of virtual network reconfiguration.

6.9 Chapter Summary

In this chapter, we proposed a novel virtual network embedding (VNE) algorithms for elastic optical networks (EONs) called alignment and consecutiveness-aware virtual network embedding (ACT-VNE) and importance, alignment and consecutiveness-aware virtual network embedding (iACT-VNE) for transparent Virtual network Embedding (t-VNE). In addition to the measure of the spectrum consecutiveness relative to the incoming virtual network (VN) requests, ACT-VNE also considers the spectrum alignment between adjacent links to calculate consecutiveness and alignment-aware node rank (ACT-Rank) which facilitates the node mapping process. iACT-VNE considers the relative importance of nodes with respect to nodes already mapped and changed the ACT node rank according to map the virtual nodes to substrate node with close proximity thereby making better use of link bandwidth. The simulation results also showed that the proposed schemes can achieve lower blocking probability for the incoming VN requests and higher revenue for the Service Providers. iACT-VNE in turn achieves less blocking probability and less average link utilization that ACT-VNE making better use of the spectrum. Moreover, the average substrate path length is also less in iACT-VNE since it considers relative importance between substrate nodes in terms of distance.

PART 7

CONCLUSIONS

The Internet data traffic is expected to grow at an unprecedented rate in the future. To handle the growth of date traffic in the future and to alleviate the issue of bandwidth crunch in the current Internet infrastructure, this dissertation mainly studied the use and the application of Elastic Optical Networks (EONs). To ease the application of EONs, this dissertation proposed schemes to solve the fundamental problems of routing and spectrum allocation (RSA) in EONs and the issue of the fragmentation of spectral resources in EONs.

The issue of spectrum fragmentation is very prominent is EONs and this issue becomes more severe with the random arrival/departure of incoming requests. For these problems, this dissertation extensively reviewed the related work and presented a classified overview of the literature study. To mitigate the issue of spectrum fragmentation, this dissertation proposed a novel spectrum defragmentation algorithms that consolidates the spectrum. This spectrum defragmentation algorithm captures the spectral fragmentation information into an auxiliary graph and then calculates the maximum independent sets to consolidate the spectrum. Simulation results confirmed that the use of this spectrum defragmentation problem can consolidate the spectrum by about 30%. While spectrum defragmentation can free up scattered spectrum, this operation can be very expensive. This dissertation also proposed multipath routing algorithm in EONs that can divide a single request of higher demand into multiple requests of lower demands and can accommodate these request through fragmented spectrum resources.

This dissertation utilized the periodic behavior of Internet data traffic for the efficient allocation of sub-carriers in EONs. To this end, this dissertation defined the problem of *time-varying spectrum allocation* in EONs where the spectrum assigned to a connection demand can expand or contract based on the traffic demand. This dissertation also formulated the

time-varying spectrum allocation problem as an Integer linear programming (ILP) formulation and also, proposed four heuristics algorithms that used the periodic behavior of Internet date traffic to estimate the future traffic demands and allocate sub-carriers to connection demands with the goal to minimize the network reconfiguration cost.

The latter part of this dissertation is dedicated to the application of EONs in network virtualization that enables the use of optical networks as a service (NaaS). With network virtualization, multiple networks can remain in isolation with each other on top of a common substrate network and the use of EONs as one of the efficient substrate technologies makes virtual network embedding more agile and scalable. To facilitate the application of EONs in network virtualization, this dissertation proposed a novel virtual network embedding algorithm that takes into account the spectrum alignment and the consecutiveness between adjacent fiber links, when mapping virtual nodes/links to substrate nodes/links. Since EONs is prone to spectrum fragmentation, this dissertation also proposed a novel virtual network reconfiguration algorithms that can reconfigure the spectrum assignment to accommodate more virtual networks and can maximize the revenue for the service providers.

REFERENCES

[1] M. Jinno, H. Takara, B. Kozicki, Y. Tsukishima, Y. Sone, and S. Matsuoka, "Spectrum-efficient and scalable elastic optical path network: architecture, benefits, and enabling technologies," *IEEE Communications Magazine*, vol. 47, no. 11, pp. 66–73, Nov. 2009.

[2] *Cisco Visual Networking Index: Forecast and Methodology, 2013-2018*, 2014. [Online]. Available: http://www.cisco.com/

[3] B. Mukherjee, *Optical Communication Networks*. Springer-Verlag, 2006.

[4] R. Ramaswami, K. Sivarajan, and G. Sasaki, *Optical Networks: A Practical Perspective, 3rd Edition*, 3rd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2009.

[5] M. Jinno, H. Takara, and B. Kozicki, "Dynamic optical mesh networks: drivers, challenges and solutions for the future," in *Proceedings of ECOC*, Sep. 2009, pp. 1–4.

[6] Y. Wang, X. Cao, and Y. Pan, "A study of the routing and spectrum allocation in spectrum-sliced elastic optical path networks," in *Proceedings of IEEE INFOCOM*, Apr. 2011, pp. 1503–1511.

[7] K. Christodoulopoulos, I. Tomkos, and E. Varvarigos, "Routing and spectrum allocation in OFDM- based optical networks," in *Proceedings of IEEE GLOBECOM*, Dec. 2010, pp. 1–6.

[8] A. Patel, P. Ji, J. Jue, and T. Wang, "Defragmentation of transparent flexible optical WDM (FWDM) networks," in *Proceedings of OFC/NFOEC*, Mar. 2011, pp. 1–3.

[9] S. Shakya and X. Cao, "Spectral defragmentation in elastic optical path networks using independent sets," in *Proceedings of OFC/NFOEC*. Optical Society of America, 2013.

[10] I. Tomkos, B. Mukherjee, S. K. Korotky, R. Tucker, and L. Lunardi, "The evolution

of optical networking," *Proceedings of the IEEE*, vol. 100, no. 5, pp. 1017–1022, May. 2012.

[11] R. Alferness, "The evolution of configurable wavelength multiplexed optical networks - a historical perspective," *Proceedings of the IEEE*, vol. 100, no. 5, pp. 1023–1034, May 2012.

[12] I. Jacobs, "Lightwave system development: Looking back and ahead," *Optical Photonic News*, vol. 6, no. 2, p. 19, Feb 1995.

[13] S. Gringeri, B. Basch, V. Shukla, R. Egorov, and T. Xia, "Flexible architectures for optical transport nodes and networks," *IEEE Communications Magazine*, vol. 48, no. 7, pp. 40–50, Jul. 2010.

[14] S. J. B. Yoo, "Optical packet and burst switching technologies for the future photonic internet," *Journal of Lightwave Technology*, vol. 24, no. 12, pp. 4468–4492, Dec 2006.

[15] A. Pattavina, "Architectures and performance of optical packet switching nodes for ip networks," *J. Lightwave Technol.*, vol. 23, no. 3, p. 1023, Mar 2005.

[16] Y. Chen, C. Qiao, and X. Yu, "Optical burst switching: a new area in optical networking research," *IEEE Networks*, vol. 18, no. 3, pp. 16–23, May 2004.

[17] X. Cao, V. Anand, and C. Qiao, "Waveband switching in optical networks," *IEEE Communications Magazine*, vol. 41, no. 4, pp. 105–112, April 2003.

[18] R. Izmailov, S. Ganguly, T. Wang, Y. Suemura, Y. Maeno, and S. Araki, "Hybrid hierarchical optical networks," *IEEE Communications Magazine*, vol. 40, no. 11, pp. 88–94, Nov 2002.

[19] B. Kozicki, H. Takara, T. Yoshimatsu, K. Yonenaga, and M. Jinno, "Filtering characteristics of highly-spectrum efficient spectrum-sliced elastic optical path (slice) network," in *Proceedings of OFC/NFOEC*, March 2009, pp. 1–3.

[20] S. Shirazipourazad, C. Zhou, Z. Derakhshandeh, and A. Sen, "On routing and spectrum allocation in spectrum-sliced optical networks," in *Proceedings of IEEE INFOCOM*, Apr. 2013, pp. 385–389.

[21] Y. Wang, X. Cao, and Q. Hu, "Routing and spectrum allocation in spectrum-sliced elastic optical path networks," in *Proceedings of IEEE ICC*, Jun. 2011, pp. 1 –5.

[22] Y. Wang, X. Cao, Q. Hu, and Y. Pan, "Towards elastic and fine-granular bandwidth allocation in spectrum-sliced optical networks," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 4, no. 11, pp. 906–917, Nov 2012.

[23] A. N. Patel, P. N. Ji, J. P. Jue, and T. Wang, "Routing, wavelength assignment, and spectrum allocation in transparent flexible optical WDM (FWDM) networks," in *Integrated Photonics Research, Silicon and Nanophotonics and Photonics in Switching*. Optical Society of America, 2010.

[24] ——, "Routing, wavelength assignment, and spectrum allocation algorithms in transparent flexible optical WDM networks," *Optical Switching and Networking*, vol. 9, no. 3, pp. 191 – 204, 2012.

[25] M. Klinkowski and K. Walkowiak, "Routing and spectrum assignment in spectrum sliced elastic optical path network," *IEEE Communications Letters*, vol. 15, no. 8, pp. 884–886, 2011.

[26] K. Christodoulopoulos, I. Tomkos, and E. Varvarigos, "Spectrally/bitrate flexible optical network planning," in *Proceedings of 36th ECOC*, Sept 2010, pp. 1–3.

[27] L. Velasco, M. Klinkowski, M. Ruiz, and J. Comellas, "Modeling the routing and spectrum allocation problem for flexgrid optical networks," *Photonic Network Communications*, vol. 24, no. 3, pp. 177–186, 2012.

[28] A. N. Patel, P. N. Ji, J. P. Jue, and T. Wang, "Dynamic routing, wavelength assignment,

and spectrum allocation in transparent flexible optical WDM networks," pp. 79 590M–79 590M–8, 2011.

[29] X. Wan, L. Wang, N. Hua, H. Zhang, and X. Zheng, "Dynamic routing and spectrum assignment in flexible optical path networks," in *Proceedings of OFC/NFOEC*, March 2011, pp. 1–3.

[30] A. Castro, L. Velasco, M. Ruiz, M. Klinkowski, J. P. FernáNdez-Palacios, and D. Careglio, "Dynamic routing and spectrum (re)allocation in future flexgrid optical networks," *Comput. Netw.*, vol. 56, no. 12, pp. 2869–2883, Aug. 2012.

[31] X. Wang, Q. Zhang, I. Kim, P. Palacharla, and M. Sekiya, "Blocking performance in dynamic flexible grid optical networks - what is the ideal spectrum granularity?" in *Proceedings of 37th ECOC*, Sept 2011, pp. 1–3.

[32] Y. Yu, J. Zhang, Y. Zhao, X. Cao, X. Lin, and W. Gu, "The first single-link exact model for performance analysis of flexible grid wdm networks," in *Proceedings of OFC/NFOEC*, March 2013, pp. 1–3.

[33] R. Almeida, R. Delgado, C. Bastos-Filho, D. Chaves, H. Pereira, and J. Martins-Filho, "An evolutionary spectrum assignment algorithm for elastic optical networks," in *Proceedings of 15th International Conference on Transparent Optical Networks (ICTON)*, June 2013, pp. 1–3.

[34] Y. Wang, J. Zhang, Y. Zhao, J. Wang, and W. Gu, "Routing and spectrum assignment by means of ant colony optimization in flexible bandwidth networks," in *Proceedings of OFC/NFOEC*, March 2012, pp. 1–3.

[35] W. Shieh, "OFDM for flexible high-speed optical networks," *Journal of Lightwave Technology*, vol. 29, no. 10, pp. 1560–1577, May 2011.

[36] H. Takara, B. Kozicki, Y. Sone, and M. Jinno, "Spectrally-efficient elastic optical

path networks," in *Proceedings of 15th Optoelectronics and Communications Conference (OECC)*, July 2010, p. 116117.

[37] B. Kozicki, H. Takara, Y. Sone, A. Watanabe, and M. Jinno, "Distance-adaptive spectrum allocation in elastic optical path network (SLICE) with bit per symbol adjustment," in *Proceedings of OFC/NFOEC*, March 2010, pp. 1–3.

[38] H. Takara, B. Kozicki, Y. Sone, T. Tanaka, A. Watanabe, A. Hirano, K. Yonenaga, and M. Jinno, "Distance-adaptive super-wavelength routing in elastic optical path network (SLICE) with optical OFDM," in *Proceedings of 36th ECOC*, Sept 2010, pp. 1–3.

[39] T. Takagi, H. Hasegawa, K. Sato, Y. Sone, B. Kozicki, A. Hirano, and M. Jinno, "Dynamic routing and frequency slot assignment for elastic optical path networks that adopt distance adaptive modulation," in *Proceedings of OFC/NFOEC*, Mar. 2011, pp. 1–3.

[40] H. Takara, K. Yonenaga, and M. Jinno, "Spectrally-efficient elastic optical path networks toward 1 tbps era," in *Proceedings of OFC/NFOEC*, March 2012, pp. 1–3.

[41] A. Nag, M. Tornatore, and B. Mukherjee, "Optical network design with mixed line rates and multiple modulation formats," *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 466–475, Feb 2010.

[42] M. Jinno, B. Kozicki, H. Takara, A. Watanabe, Y. Sone, T. Tanaka, and A. Hirano, "Distance-adaptive spectrum resource allocation in spectrum-sliced elastic optical path network," *IEEE Communications Magazine*, vol. 48, no. 8, pp. 138–145, August 2010.

[43] A. Morea and O. Rival, "Advantages of elasticity versus fixed data-rate schemes for restorable optical networks," in *Proceedings of 36th ECOC*, Sept 2010, pp. 1–3.

[44] S. Yang and F. Kuipers, "Impairment-aware routing in translucent spectrum-sliced elastic optical path networks," in *Proceedings of 17th European Conference on Networks and Optical Communications (NOC)*, June 2012, pp. 1–6.

[45] K. Christodoulopoulos, I. Tomkos, and E. Varvarigos, "Elastic bandwidth allocation in flexible OFDM-based optical networks," *Journal of Lightwave Technology*, vol. 29, no. 9, pp. 1354–1366, May 2011.

[46] H. Beyranvand and J. Salehi, "A quality-of-transmission aware dynamic routing and spectrum assignment scheme for future elastic optical networks," *Journal of Lightwave Technology*, vol. 31, no. 18, pp. 3043–3054, Sept 2013.

[47] R. Wang and B. Mukherjee, "Spectrum management in heterogeneous bandwidth optical networks," *Optical Switching and Networking,*, vol. 11, Part A, no. 0, pp. 83 – 91, 2014.

[48] N. Kadu, S. Shakya, and X. Cao, "Modulation-aware multipath routing and spectrum allocation in elastic optical networks," in *Proceedings of IEEE International Conference on Advanced Networks and Telecommuncations Systems (ANTS)*, Dec 2014, pp. 1–6.

[49] N. Amaya, M. Irfan, G. Zervas, K. Banias, M. Garrich, I. Henning, D. Simeonidou, Y. Zhou, A. Lord, K. Smith, V. Rancano, S. Liu, P. Petropoulos, and D. Richardson, "Gridless optical networking field trial: Flexible spectrum switching, defragmentation and transport of 10G/40G/100G/555G over 620-km field fiber," in *Proceedings of 37th ECOC*, Sept 2011, pp. 1–3.

[50] Z. Zhu, W. Lu, L. Zhang, and N. Ansari, "Dynamic service provisioning in elastic optical networks with hybrid single-/multi-path routing," *Journal of Lightwave Technology*, vol. 31, no. 1, pp. 15–22, Jan 2013.

[51] M. Zhang, W. Lu, Z. Zhu, Y. Yin, and S. Yoo, "Planning and provisioning of elastic O-OFDM networks with fragmentation-aware routing and spectrum assignment (rsa) algorithms," in *Communications and Photonics Conference (ACP)*, Nov 2012, pp. 1–3.

[52] X. Wang, Q. Zhang, I. Kim, P. Palacharla, and M. Sekiya, "Utilization entropy for assessing resource fragmentation in optical networks," in *Proceedings of OFC/NFOEC*, March 2012, pp. 1–3.

[53] K. Christodoulopoulos, I. Tomkos, and E. Varvarigos, "Dynamic bandwidth allocation in flexible OFDM-based networks," in *Proceedings of OFC/NFOEC*, 2011, pp. 1–3.

[54] A. Castro, L. Velasco, M. Ruiz, M. Klinkowski, J. P. FernáNdez-Palacios, and D. Careglio, "Dynamic routing and spectrum (re)allocation in future flexgrid optical networks," *Journal of Computer Networks*, vol. 56, no. 12, pp. 2869–2883, Aug. 2012.

[55] F. Cugini, F. Paolucci, G. Meloni, G. Berrettini, M. Secondini, F. Fresi, N. Sambo, L. Poti, and P. Castoldi, "Push-pull defragmentation without traffic disruption in flexible grid optical networks," *Journal of Lightwave Technology*, vol. 31, no. 1, pp. 125–133, Jan 2013.

[56] K. Wen, Y. Yin, D. Geisler, S. Chang, and S. Yoo, "Dynamic on-demand lightpath provisioning using spectral defragmentation in flexible bandwidth networks," in *Proceedings of 37th ECOC*, Sept 2011, pp. 1–3.

[57] R. Proietti, C. Qin, B. Guan, Y. Yin, R. P. Scott, R. Yu, and S. J. B. Yoo, "Rapid and complete hitless defragmentation method using a coherent rx lo with fast wavelength tracking in elastic optical networks," *Opt. Express*, vol. 20, no. 24, pp. 26 958–26 968, Nov 2012.

[58] X. Chen, Y. Zhong, and A. Jukan, "Multipath routing in elastic optical networks with distance-adaptive modulation formats," in *Proceedings of IEEE ICC*, June 2013, pp. 3915–3920.

[59] W. Lu, X. Zhou, L. Gong, M. Zhang, and Z. Zhu, "Dynamic multi-path service provisioning under differential delay constraint in elastic optical networks," *IEEE Communications Letters*, vol. 17, no. 1, pp. 158–161, January 2013.

[60] A. Castro, L. Velasco, M. Ruiz, and J. Comellas, "Single-path provisioning with multipath recovery in flexgrid optical networks," in *Proceedings of 4th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Oct 2012, pp. 745–751.

[61] L. Ruan and N. Xiao, "Survivable multipath routing and spectrum allocation in OFDM-based flexible optical networks," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 5, no. 3, pp. 172–182, March 2013.

[62] X. Chen, A. Jukan, and A. Gumaste, "Optimized parallel transmission in elastic optical networks to support high-speed ethernet," *Journal of Lightwave Technology*, vol. 32, no. 2, pp. 228–238, Jan 2014.

[63] L. Ruan and Y. Zheng, "Dynamic survivable multipath routing and spectrum allocation in OFDM-based flexible optical networks," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 6, no. 1, pp. 77–85, Jan 2014.

[64] G. Shen, Q. Yang, S. You, and W. Shao, "Maximizing time-dependent spectrum sharing between neighbouring channels in CO-OFDM optical networks," in *Proceedings of 13th International Conference on Transparent Optical Networks (ICTON)*, 2011, pp. 1–4.

[65] J. Zhang, J. Zhang, Y. Zhao, and X. Wang, "Time-dependent spectrum resource sharing in flexible bandwidth optical networks," *IET Networks*, vol. 1, no. 4, pp. 189–198, 2012.

[66] C. Xie, J. Zhang, Y. Zhao, J. Zhang, H. Yang, and X. Yu, "Spectrum sharing for time-varying traffic in openflow-based flexi-grid optical networks," in *Communications and Photonics Conference (ACP)*, 2012, pp. 1–3.

[67] E. Palkopoulou, I. Stiakogiannakis, D. Klonidis, K. Christodoulopoulos, E. Varvarigos, O. Gerstel, and I. Tomkos, "Dynamic cooperative spectrum sharing in elastic networks," in *Proceedings of OFC/NFOEC*. Optical Society of America, 2013.

[68] M. Klinkowski, M. Ruiz, L. Velasco, D. Careglio, V. Lopez, and J. Comellas, "Elastic spectrum allocation for time-varying traffic in flexgrid optical networks," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 1, pp. 26–38, 2013.

[69] Z. Ye, X. Cao, X. Gao, and C. Qiao, "Predictive and incremental grooming scheme for

time-varying traffic in wdm network," in *Proceedings of INFOCOM' 13 Mini-conference*, 2013.

[70] S. Shakya, X. Cao, Z. Ye, and C. Qiao, "Spectrum allocation in spectrum-sliced elastic optical path networks using traffic prediction," *Photonic Network Communications*, pp. 1–12, 2015.

[71] S. Shakya, Y. Wang, X. Cao, Z. Ye, and C. Qiao, "Minimize sub-carrier reallocation in elastic optical path networks using traffic prediction," in *Proceeding of IEEE GLOBECOM*, Dec 2013, pp. 2352–2357.

[72] S. Shakya, X. Cao, Z. Ye, and C. Qiao, "Spectrum allocation for time-varying traffic in elastic optical networks using traffic pattern," in *Proceedings of OFC/NFOEC*. Optical Society of America, 2014.

[73] M. F. Zhani and H. Elbiaze, "Analysis and prediction of real network traffic," *Journal of Networks*, vol. 4, no. 9, pp. 855–865, Dec. 2009.

[74] Y. Chen, S. Jain, V. Adhikari, Z.-L. Zhang, and K. Xu, "A first look at inter-data center traffic characteristics via yahoo! datasets," in *Proceedings of IEEE INFOCOM*, Apr. 2011, pp. 1620–1628.

[75] S. Sundaresan, W. de Donato, N. Feamster, R. Teixeira, S. Crawford, and A. Pescapè, "Broadband internet performance: a view from the gateway," *SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 4, pp. 134–145, Aug. 2011.

[76] I. Chlamtac, A. Ganz, and G. Karmi, "Lightpath communications: an approach to high bandwidth optical wan's," *IEEE Transactions on Communications*, vol. 40, no. 7, pp. 1171–1182, Jul. 1992.

[77] A. Patel, C. Gao, J. Jue, X. Wang, Q. Zhang, P. Palacharla, and T. Naito, "Traffic grooming and regenerator placement in impairment-aware optical WDM networks," in *Proceedings of 14th ONDM*, Feb 2010, pp. 1–6.

[78] R. M. Karp and A. Wigderson, "A fast parallel algorithm for the maximal independent set problem," *Journal of the ACM*, vol. 32, no. 4, pp. 762–773, Oct. 1985.

[79] A. Sharieh and W. Rawagepfeh, "An algorithm for finding maximum independent set in a graph," *European Journal of Science*, vol. 23, pp. 586–596, 2008.

[80] H. Tong, C. Li, J. He, and Y. Chen, "Internet traffic prediction by w-boost: Classification and regression," vol. 3498, pp. 397–402, 2005.

[81] I. CPLEX. [Online]. Available: http://www.ilog.com

[82] N. Feamster, L. Gao, and J. Rexford, "How to lease the internet in your spare time," *SIGCOMM Comput. Commun. Rev.*, vol. 37, no. 1, pp. 61–64, Jan. 2007.

[83] M. Yu, Y. Yi, J. Rexford, and M. Chiang, "Rethinking virtual network embedding: Substrate support for path splitting and migration," *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 2, pp. 17–29, Mar. 2008.

[84] A. Pages, J. Perello, S. Spadaro, J. Garcia-Espin, J. Riera, and S. Figuerola, "Optimal allocation of virtual optical networks for the future internet," in *Proceedings of 16th International Conference on Optical Network Design and Modeling (ONDM),*, Apr 2012, pp. 1–6.

[85] L. Gong and Z. Zhu, "Virtual optical network embedding (VONE) over elastic optical networks," *IEEE Journal of Lightwave Technology*, vol. 32, no. 3, pp. 450–460, Feb 2014.

[86] A. Pages, J. Perello, S. Spadaro, J. Garcia-Espin, J. Riera, and S. Figuerola, "Optimal allocation of virtual optical networks for the future internet," in *16th International Conference on Optical Network Design and Modeling (ONDM)*, Apr. 2012, pp. 1–6.

[87] J. Zhao, S. Subramaniam, and M. Brandt-Pearce, "Virtual topology mapping in elastic optical networks," in *Proceedings of IEEE International Conference on Communications (ICC)*, June 2013, pp. 3904–3908.

[88] Y. Sone, A. Hirano, A. Kadohata, M. Jinno, and O. Ishida, "Routing and spectrum assignment algorithm maximizes spectrum utilization in optical networks," in *Proceedings of 37th ECOC*, Sep. 2011, pp. 1–3.

[89] D. Andersen, "Theoretical approaches to node assignment," in *unpublished Manuscript*, Dec. 2002.

[90] S. White and P. Smyth, "Algorithms for estimating relative importance in networks," in *Proceedings of the Ninth ACM SIGKDD*, 2003, pp. 266–275.

[91] P. Tran and A. Timm-Giel, "Reconfiguration of virtual network mapping considering service disruption," in *Proceedings of IEEE International Conference on Communications (ICC)*, Jun. 2013, pp. 3487–3492.

[92] Y. Zhu and M. Ammar, "Algorithms for assigning substrate network resources to virtual network components," in *Proceedings of IEEE INFOCOM*, Apr. 2006, pp. 1–12.

[93] N. Farooq Butt, M. Chowdhury, and R. Boutaba, "Topology-awareness and reoptimization mechanism for virtual network embedding," in *Proceedings of the 9th IFIP TC 6 International Conference on Networking*, ser. NETWORKING'10. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 27–39.

[94] I. Fajjari, N. Aitsaadi, G. Pujolle, and H. Zimmermann, "VNR algorithm: A greedy approach for virtual networks reconfigurations," in *Proceedings of IEEE GLOBECOM*, Dec. 2011, pp. 1–6.

[95] F. Gu, M. Peng, S. Khan, A. Rayes, and N. Ghani, "Virtual network reconfiguration in optical substrate networks," in *Proceedings of OFC/NFOEC*, Sep. 2013, p. NTh4J.6.

[96] N. A. Boutheina Dab, Ilhem Fajjariz and G. Pujolle, "VNR-GA: Elastic virtual network reconfiguration algorithm based on genetic metaheuristic," in *Proceedings of the IEEE GLOBECOM*, 2013.

[97] S. Shakya, N. Pradhan, X. Cao, Z. Ye, and C. Qiao, "Virtual network embedding and reconfiguration in elastic optical networks," in *Proceedings of GLOBECOM*, Dec 2014, pp. 2160–2165.