Optimal Placement of Reconfigurable Optical Add/Drop Multiplexers with Packing, Blocking, and Signal Loss

Conrado Borraz-Sánchez[1], Diego Klabjan
{c-borraz-sanchez, d-klabjan}@northwestern.edu

Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL 60208

Abstract

With technological and manufacturing advances, and increased economies of scale, today the use of Reconfigurable Optical Add/Drop Multiplexers (ROADMs) has become economical. A ROADM allows individual or multiple wavelengths to be added and/or dropped from a transport fiber without the need to convert the signals to electrical and back again to optical. This feature makes a fiber optic network more flexible, i.e., the network can be remotely configured/reconfigured through ROADMs. Consequently, due to high capital cost, to determine an optimal placement of ROADMs and to assign wavelengths in a cost efficient way becomes an important problem in network topology design. We introduce a Mixed Integer Programming model that captures signal loss and wavelength packing/blocking, and develop a three-phase algorithm to efficiently solve it. Analytical justifications for the algorithmic steps are provided. Computational experiments are conducted to assess the tractability of the model, and to evaluate the performance of the proposed approach. The execution times of the algorithm are acceptable.

Keywords: Add/Drop Multiplexer Placement, ROADM, WDM Network, Mixed-Integer Programming, Wavelength packing/blocking
2000 MSC: 9004, 90C08, 94C04

1. Introduction

Optimal Wavelength-Division Multiplexing (WDM) is a promising technology for the telecommunication industry [1]. The technology has advanced in recent years to the point of becoming financially feasible. Optical signal transport is substantially cheaper than the transfer of an electrical signal. As a result, it becomes more economical to transfer signals over optic cables and thus to minimize the transport of electrical signals, which are required at end points.

[1]Corresponding author

In a fiber optic network, an electrical signal is transformed into an optical signal and vice versa on several occasions. Every service using the network sends an electrical signal that must be converted via transponders and muxponders to an optical signal for transport. In addition, transponders and muxponders also prevent the converted optical signals from exceeding the capacity of a wavelength (typically 10G). The converted optical signal is thus transferred through the network to its final destination at which an optical-to-electrical conversion process is applied.

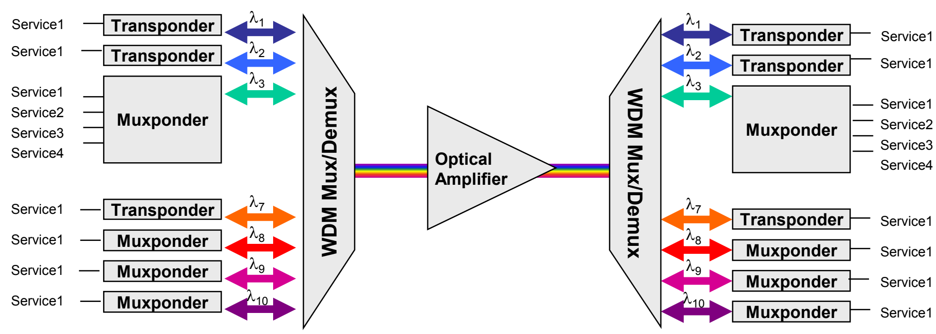

Figure 1: Network transfer on a single line

Figure 1 shows a signal conversion process when applied on a single line. A wavelength is a bundle of services with a unique frequency (labeled $\lambda_1, \ldots, \lambda_{10}$ in Figure 1). Each wavelength is packed into the fiber optic cable and in essence travels independently of the remaining wavelengths.

During the transmission process, the optical signals must be amplified every certain distance along a fiber optic cable (typically within a range of 100 miles). The amplification process, however, causes a degradation in signal quality (noisy increase). To overcome this, the signal is regenerated periodically. Figure 1 shows a single amplifier.

A Reconfigurable Optical Add/Drop Multiplexer (ROADM) can automatically terminate wavelengths from one optical cable to another (see Figure 2). Wavelengths can be routed across the network at the optical layer or "dropped" at traffic access points to provide service. In addition, a ROADM can terminate a wavelength for a group of services, which corresponds to converting the optical signal through transponders and muxponders back to electrical signals used by the services. This operation is called a service add/drop.

Wavelength termination includes wavelength repacking or blocking. Within a ROADM, individual wavelengths are de-multiplexed and then switched by an optical switch module. Any wavelength can be directed to any output port. After switching, the optical signals are then repacked (i.e., all outgoing transmitted data are grouped/multiplexed) into output fibers. As shown on the left in Figure 2, the data carried in wavelength c_1 is terminated and/or repacked by a line card, and then the resulting wave packets are sent out in wavelengths c_2 and c_3.

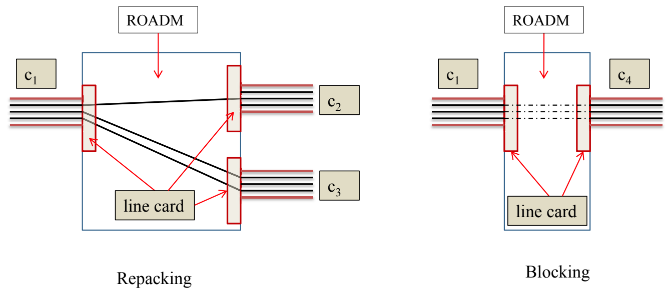

Figure 2: Wavelength termination process

Figure 2 also shows an example of wavelength blocking. On the right, wavelength c_1 is converted from optical to electric by a line card. This is required when output links do not match an identical incoming wavelength. Hence, c_1 is blocked and the wave packet while converted back to optical by the line card is sent out as wavelength c_4. The actual repacking and blocking is performed at the electric level. A ROADM is required if a wavelength is blocked, or repacked, or there is a service add/drop at a site. If any of these operations is performed, the underlying wavelength must be terminated.

In fiber optic networks, wavelength blocking has a key role for a successful completion of a signal transfer. This stems from the fact that an optic cable can include only wavelengths of different frequencies, and in the entire system there is only a finite number of available frequencies. When the frequency of a wavelength corresponds to the same frequency of another wavelength in the system, the blocking process is performed by ROADMs and the frequency of a wavelength must be changed.

Termination of a wavelength is required only when proceeding with a wavelength other than an identical wavelength (i.e., same frequency and composition of services). Once the wavelength is terminated, the cost of the line cards is incurred, coupled with the cost of a ROADM (which must be installed if there is at least one termination).

The downside of using ROADMs is the weakening of the signal and the occurrence of additional noise. To overcome these, amplifiers are installed every certain distance along the line to amplify signals (wavelengths). Figure 3 shows the locations of three amplifiers, which are predetermined based on the network configuration. The network segment between two consecutive amplifiers is considered the smallest segment where no changes to the cable and signals are possible.

In addition to signal loss, signals also degrade in quality and must be regenerated. To regenerate a signal, a line card is required, i.e., the signal must be converted to electrical, regenerated, and then converted back to optical. The need for regeneration depends on the number of segments between two locations, the number of ROADMs in between (each ROADM further degrades the signal), and the signal loss factor of the underlying cable. As a result, a number of regenerators must be deployed in the network.

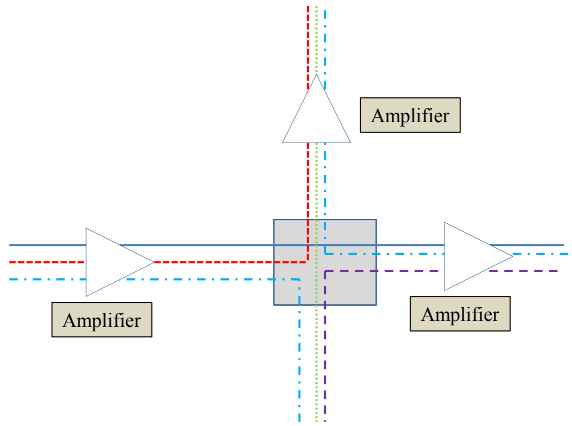

Figure 3: Amplification of signal

Due to the high cost of ROADMs and line cards observed in the past on existing networks, ROADMs were not economically feasible and thus they were not deployed. Large telecommunication companies using fiber optic networks without ROADMs had to route all their services directly.

The telecommunication industry has evolved and advanced. With technological and manufacturing advances, coupled with the increased economies of scale, today ROADMs have become economically feasible. As a consequence, providers are upgrading their high-bandwidth networks to include ROADMs in order to improve efficiency and flexibility. Due to the substantial increase in the use of ROAMDs on WDM optical networks, only a small marginal improvement in the costs of installing and using ROADMs yields a substantial economic impact. On the other hand, optical network systems including many wavelengths and operating at a high data rate lead to a significant amount of data loss, thus affecting the overall quality of service. Therefore, ROADMs' optimal use and accurate measurement in terms of optical signal-to-noise ratio became key metrics of success in WDM optical network applications.

In this paper we focus on the design of a fiber optic network with ROADMs in the most cost effective way, and confine our study to mesh networks. We assume that the existing cable network and the underlying services are given. A service is defined by two end points and a route or path consisting of a number of segments connecting the two end points. A service is mapped to a set of wavelengths based on the number of circuits required. Three different circuits are given, namely, 10G, 2.5G and 1G circuits. The main goal of this study is to develop an algorithm to solve the splittable traffic partition in WDM optical networks. The objective is to develop an algorithm capable of: (1) minimizing the total cost including the cost of ROADMs and line cards, (2) obeying all engineering requirements, e.g., all services packed into a wavelength on each segment should not exceed the maximum capacity of a wavelength, and (3) computationally handling large-scale instances arising in the industry.

1.1. Related Work

There have already been studies of placement of ROADMs. For example, Sutter et al. [2], Belvaux et al. [3], and Lee and Koh [4] consider a network connected by unidirectional Synchronous Optical Networking (SONET) circular optic fibers, called rings, and propose optimization methods, such as integer programming column generation, to find the optimal placement of ROADMs. A major difference with the present work is the fact that the authors neglect wavelengths in their model. Fortz et al. [5] study a SONET ring design problem with the objective of finding a minimum cost assignment of ROADMs of unidirectional rings subject to a limitation on the number of nodes allowed in a ring, and hop constraints, i.e., constraints forcing each ring to have at least two different interconnection nodes with other rings. Tabu search algorithms are designed to assign traffic demand to wavelengths. Besides the difference of wavelength continuity, which is included in our model but disregarded in [5], our work does not include hop constraints since we have a general network topology.

Goldschmidt et al. [6] present several greedy algorithms (edge-, cut-, and node-based) to minimize the total cost of a network connected by bidirectional SONET rings. Macambira et al. [7] extend the work of Goldschmidt et al. [6] and present a branch-and-price algorithm. When compared to our work, two major differences are found in the works of Goldschmidt et al. and Macambira et al.: 1) the objective minimized in these works is the number and length of rings instead of minimizing the cost of ROADMs, and 2) their work does not capture packing, blocking, and signal loss.

All of the existing algorithms (except for Calinescu and Wan [8], Gerstel et al. [9], and Holler and Voβ [10]) assume wavelength-continuity, i.e., the same wavelength is allocated on all of the links in the path established for a service. Since a ROADM allows wavelength-discontinuity, i.e., services can be transferred from one wavelength to another, Gerstel et al. [9] argued that the cost of the network, or the number of ROADMs, can be potentially reduced by adopting this feature. Hence, minimizing the total cost of a SONET/WDM network is equivalent to finding an optimal traffic partition that uses the minimum number of ROADMs since the cost of ROADMs are the dominant cost factor. This problem is proven to be NP-hard in [8].

Moreover, Gerstel et al. [9] provide two heuristic algorithms (cut- and assign-first), and Calinescu and Wan [8] present heuristic approximation algorithms to find a lower bound of the number of ROADMs. Unlike our work, the detailed wavelength termination (packing and blocking) is ignored in their models.

Holler and Voβ [10] consider a large network and employ a two-layer approach. The authors propose greedy heuristics to find different routings and minimal total cost of ROADMs and other equipments. As previous works, wavelength termination (packing and blocking) is also ignored in [10], which is an important simplification and deviation from our model.

Hongyue and Biswanath [11] investigate switching- and tuning-based architectures using ROADMs and propose several heuristics for solving the dynamic traffic provisioning problem. They divide the problem into two subproblems: the tuning-head positioning problem and the resource allocation problem. They conclude that adding more ROADMs at the nodes is the best way to overcome the tuning constraints, yet they keep open the question of what the optimal number of ROADMs is.

More recently, Fourneau et al. [12] optimize the design of ROADM-based networks for given routing matrices and propose an analytical formula to obtain the number of packets waiting in the buffer. They basically deal with possible modifications of existing optical add/drop multiplexers to improve the network using a partial optical-electrical-optical conversion. Unlike our work, the authors mainly focus on developing a tool to improve the performance of ROADMs to design core networks with aggregation channels, and base their study on a stochastic model and greedy algorithms. Wei et al. [13] present a two-way hitless ROADM structure for a two-way network, and an Optical Cross Connect structure for a mesh network. Their work is based on an add/drop filter, called microring resonator (MRR), to access one channel in WDM signals while the other channels are not being disturbed. Unlike the present work, the authors focus on a combined ROADM-MRR structure to optimize the port number using the structure as node module and adjusting the index of MRR.

The main contributions of our work are as follows:

- We capture signal degradation and loss. It is the first work addressing signal loss in the content of placement of ROADMs.
- We address the wavelength-discontinuity nature offered by ROADMs and line cards, and incorporate packing and blocking into our model directly. In contrast, prior research considering wavelength-discontinuity ([8], [9], [10]) assume wavelength is discontinuous, but ignore the detailed implementation of packing and blocking in their models. The allowance of packing and blocking poses a significant challenge.
- We consider a general network topology. In contrast, existing models are based on networks with rings, or the origin and destination of services are the same.
- We design and implement a three-phase optimization algorithm to effectively tackle the ROADM placement and wavelength assignment minimization problem in WDM optical networks. As the main core of the proposed approach we introduce a novel mixed-integer programming model that efficiently incorporates wavelength termination.

The remainder of this paper is organized as follows. In the next section, we present the specifics of the problem, and introduce the mixed-integer-programming model. In Section 3, we propose a three-phase heuristic algorithm to solve the problem efficiently. Section 4 provides a computational study on the proposed heuristic. Finally, concluding remarks are given in Section 5.

2. Mathematical Formulation

In this section, we introduce a mathematical formulation of the problem with the aim of efficiently incorporating and evaluating wavelength termination. The formulation is given as a mixed integer linear program, and employed as the main core of the algorithm proposed in the next section. Note that all information concerning the underlying network and services are provided in advance, as well as all routes of traffic of a service involved in the network.

2.1. Definitions and Notations

Given a set of nodes N, a graph $G = (N, E)$ connects these nodes via a set of edges (fiber optic segments) E. Nodes correspond to locations with service requirements, sites with amplifiers, and sites where traffic can be combined. An edge corresponds to a segment as discussed in the introduction, i.e., the smallest portion of the physical network not requiring any change to the signal.

Given edge $e = (i, j) \in E$, we denote by $I(e) = \{i, j\}$ the set of the two nodes incident to the edge. It follows that $I(e) \cap I(f) \neq \emptyset$ if and only if edges e and f share the same node.

A service is specified by a route, the number of requested circuits, and the capacity of each circuit. A circuit is the smallest indivisible unit of a service segment. A wavelength consists of several circuits. For example, a service might include 3 circuits of 10G capacity and a different service possibly from the same customer might require 5 circuits of capacity 2.5G. Let D denote the set of all services. For $d \in D$, let d^P be the route specified as the set of edges, and let d^S be the two terminal nodes of service d. Let v_d be the capacity of service d, and D_d the number of circuits in service d. We define C as the set of all possible wavelengths, and U as the capacity per wavelength.

Figure 4 depicts some notation and definitions of decision variables discussed below. Node i has two incident edges e and f, edge e has two wavelengths c_1 and c_2, edge f has four wavelengths labeled from c_2 through c_5. The solid black lines indicate different circuits. Line cards are used to repack or block wavelengths. For example, in Figure 4, wavelength c_1 on edge e is terminated by a line card, and new wavelengths c_3, c_4 are repacked. We recall that wavelength c on edge $e \in E$ at node $i \in N, i \in I(e)$ is not terminated only if there exists another edge f with $I(e) \cap I(f) \neq \emptyset$ such that all circuits in wavelength c on edge f are the same as all circuits in wavelength c on edge e. For example, in Figure 4, wavelength c_2 is not terminated. Wavelength c_5 on edge f is an add/drop point, i.e., it provides access to services. Likewise, the circuits in wavelengths c_3 and c_4 that are not "extended" to edge e provide add/drop access.

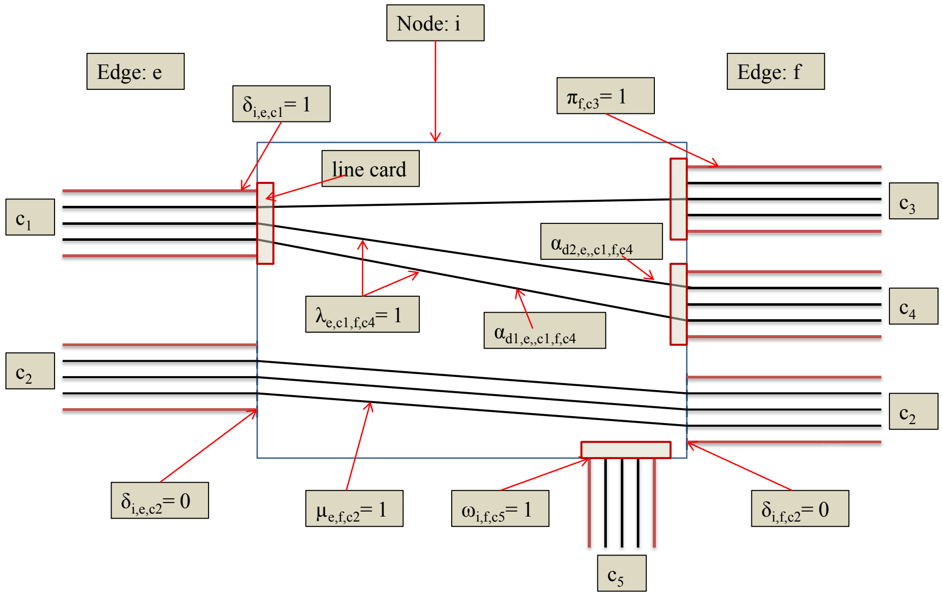

Figure 4: Parameters depiction of the mathematical model

2.2. Decision Variables

In order to model the desired features, we must capture the composition of each wavelength on each edge. In addition, for termination, we must also capture transfers among wavelengths.

- We define $\beta_{d,e,c}$ as the number of circuits of service d we allocate to wavelength c on edge e for $d \in D, c \in C$ and $e \in d^P$. For example, in Figure 4, β_{d,e,c_1} represents the number of circuits of service d allocated to wavelength c_1 on edge e.

- We define $\alpha_{d,e,c;f,\bar{c}}$ to be the number of circuits of service d going from wavelength c on edge e to wavelength $\bar{c}$ on edge f for $d \in D, c \in C, \bar{c} \in C, e \in d^P, f \in d^P, I(e) \cap I(f) \neq \emptyset, e \neq f$. As shown in Figure 4, service d_1 on wavelength c_1 of edge e transfers $\alpha_{d_1,e,c_1;f,c_4}$ circuits to edge f on wavelength c_4.

- For every $i \in N$, we define

$$\gamma_i = \begin{cases} 1 & \text{if ROADM is required at node } i, \\ 0 & \text{otherwise.} \end{cases}$$

- For each $c \in C$, edge $e \in E$, and node $i \in N, i \in I(e)$, we define

$$\delta_{i,e,c} = \begin{cases} 1 & \text{if wavelength } c \in C \text{ on edge } e \text{ is terminated at node } i, \\ 0 & \text{otherwise.} \end{cases}$$

  For example, in Figure 4, δ_{i,e,c_1} is 1 since wavelength c_1 is terminated. In addition, $\delta_{i,f,c_3} = 1$ and $\delta_{i,f,c_4} = 1$. But $\delta_{i,e,c_2} = 0$ since all circuits within wavelength c_2 are transferred to edge f within the same wavelength c_2 without termination.

- For each $c \in C$, edges $e \in E, f \in E, e \neq f$ with $I(e) \cap I(f) \neq \emptyset$, we define

$$\mu_{e,f,c} = \begin{cases} 0 & \text{if wavelength } c \text{ on edges } e \text{ and } f \text{ is the same, i.e.,} \\ & \text{the wavelength does not have to be terminated,} \\ 1 & \text{otherwise.} \end{cases}$$

In Figure 4, μ_{e,f,c_2} and μ_{f,e,c_2} are 1 since wavelength c_2 is not terminated.

- For each $c \in C$, edge $e \in E$, and node $i \in N, i \in I(e)$, we define

$$\omega_{i,e,c} = \begin{cases} 1 & \text{if wavelength } c \in C \text{ on edge } e \text{ includes a circuit} \\ & \text{starting/ending at node } i, \\ 0 & \text{otherwise.} \end{cases}$$

For example, in Figure 4, $\omega_{i,f,c_5} = 1$ since wavelength c_5 starts/finishes at node i and edge f. We also have $\omega_{i,f,c_3} = \omega_{i,f,c_4} = 1$.

- For each $e \in E, c \in C$, we define

$$\pi_{e,c} = \begin{cases} 1 & \text{if wavelength } c \text{ is used on edge } e, \\ 0 & \text{otherwise.} \end{cases}$$

For example, in Figure 4, π_{f,c_3} is 1 since wavelength c_3 is used on edge f. Indeed, Figure 4 depicts only wavelengths with $\pi = 1$ (all others are not shown).

- For $e \in E, f \in E, e \neq f$ with $I(e) \cap I(f) \neq \emptyset$, and $c \in C, \bar{c} \in C$, we define

$$\lambda_{e,c;f,\bar{c}} = \begin{cases} 1 & \text{if there is a circuit in wavelength } c \in C \text{ on edge } e \text{ that} \\ & \text{connects to wavelength } \bar{c} \in C \text{ on edge } f, \\ 0 & \text{otherwise.} \end{cases}$$

For example, in Figure 4, $\lambda_{e,c_1;f,c_4}$ is 1 since wavelength c_1 on edge e is connected to wavelength c_4 on edge f.

- For each path $P, c \in C$, we define

$$\zeta_{P,c} = \begin{cases} 1 & \text{if wavelength } c \text{ is used on the entire path } P, \\ 0 & \text{otherwise.} \end{cases}$$

- For each path P, and $1 \leq k \leq |P| - 2$, we define

$$\theta_{P,k} = \begin{cases} 1 & \text{if there are more than or equal to } k \text{ ROADMs on path } P, \\ 0 & \text{otherwise.} \end{cases}$$

Variables ζ and θ are required to capture signal loss, which as discussed depends on the number of ROADMs in path P.

2.3. Objective Function

The objective function reads

$$\min \sum_{\substack{i\in N, e\in E, i\in I(e)\\ c\in C}} CT_i(\delta_{i,e,c} + \omega_{i,e,c}) + \sum_{i\in N} CR_i\gamma_i,$$

where CR_i is the cost of a ROADM at node $i \in N$, and CT_i is the cost of terminating a wavelength at node $i \in N$ (the cost of a line card). It minimizes the total system cost consisting of the cost of ROADMs and line cards.

2.4. Assignment and Transfer Constraints

We first include a set of constraints that properly assigns each network component to the system while taking into consideration the corresponding physical assumptions.

$$\sum_{c\in C} \beta_{d,e,c} = D_d, \quad \forall\, d \in D, e \in d^P, I(e) \cap d^S \neq \emptyset, \tag{1}$$

$$\sum_{c\in C} \beta_{d,e,c} = \sum_{c\in C} \beta_{d,f,c}, \quad \forall\, d \in D, e \in d^P, f \in d^P, I(e) \cap I(f) \neq \emptyset, \tag{2}$$

$$\sum_{d\in D} v_d\beta_{d,e,c} \leq U, \quad \forall\, e \in E, c \in C. \tag{3}$$

Eq. (1) forces all circuits from each service to be assigned to the fiber optic network. Eq. (2) states the network flow balance at each "internal" node of a route, i.e., all circuits entering a node must also leave the node. Eq. (3) imposes an upper bound U on the maximum capacity of a wavelength.

We must link the content within a wavelength (variable β) with the variable (α) capturing transfers within a node. Let

$$\Omega = \{(d, f, \bar{c}, e, c) | d \in D, c \in C, \bar{c} \in C, e \in d^P, f \in d^P, I(e) \cap I(f) \neq \emptyset, e \neq f.\}$$

be the feasible search space for the signal transfer design part. Then, the following set of constraints capture signal transfers in the system.

$$\begin{aligned}
\beta_{d,e,c} &= \sum_{\bar{c}\in C} \alpha_{d,e,c;f,\bar{c}}, && \forall\, (d, e, c, f, \bar{c}) \in \Omega, && (4)\\
\alpha_{d,e,c,f,\bar{c}} &= \alpha_{d,f,\bar{c},e,c}, && \forall\, (d, e, c, f, \bar{c}) \in \Omega, && (5)\\
\alpha_{d,e,c;f,\bar{c}} &\leq m_d\lambda_{e,c;f,\bar{c}}, && \forall\, (d, e, c, f, \bar{c}) \in \Omega, && (6)\\
\beta_{d,e,c} &\leq U\omega_{i,e,c}, && \forall\, d \in D, c \in C, e \in E, i \in I(e), i \in I(e) \cap d^S. && (7)
\end{aligned}$$

For any node, Eq. (4) states the transfer relationship between the flow-in circuits and a wavelength. Since the edges are bidirectional, Eq. (5) forces the transfer variables to be symmetric. Eq. (6) establishes that if a transfer within a node occurs for at least one service, then the corresponding λ must be 1. Here $m_d = \left\lfloor \frac{U}{v_d} \right\rfloor, d \in D$ is an upper bound on the number of circuits that can be packed into a wavelength. Analogously, Eq. (7) establishes that if a new circuit is added at a node, then the corresponding ω must be 1.

2.5. Wavelength Termination Constraints

The following set of constraints capture wavelengths' blocking and packing features.

Let $|C|$ be the number of wavelengths, and let $deg(i)$ be the degree of node i. Eq. (8) then imposes the termination required at edge connect $e \in E, f \in E, I(e) \cap I(f) \neq \emptyset$ on wavelength c, with $l_i = |C| \times deg(i)$.

$$\omega_{i,e,c} + \omega_{i,f,c} + \sum_{\substack{h,h\neq e,h\neq f,\bar{c}\in C \\ i\in I(h)\cap I(e)}} \lambda_{e,c,h,\bar{c}} + \sum_{\substack{h,h\neq e,\bar{c}\in C \\ h\neq f i\in I(h)\cap I(e)}} \lambda_{f,c,h,\bar{c}} + \sum_{\bar{c}\in C,\bar{c}\neq c} (\lambda_{e,c,f,\bar{c}} + \lambda_{e,\bar{c},f,c}) \\ \leq (2+2l_i)\mu_{e,f,c}, \quad \forall\, c \in C, e \in E, f \in E, i \in I(e) \cap I(f). \tag{8}$$

Termination is needed at edge connect $e \in E, f \in E, I(e) \cap I(f) \neq \emptyset$ on wavelength c if

(a) a new circuit is added on wavelength c on either edge e or f, or
(b) wavelength c on edge e has a circuit to transfer to a different wavelength on edge $h, h \neq e, h \neq f$, or
(c) same as (b) except replace e with f, or
(d) wavelength c on edge e or f transfers to a wavelength $\bar{c}, \bar{c} \neq c$ on edge f or e, respectively.

Any one of the above cases must imply $\mu_{e,f,c} = 1$. In Eq. (8), the first two terms capture case (a), the next two terms cases (b) and (c), and the last term case (d).

$$\mu_{e,f,c} \leq l_i \delta_{i,e,c}, \quad \forall\, e \in E, f \in E, i \in I(e) \cap I(f), c \in C, \tag{9}$$

$$\mu_{e,f,c} \leq l_i \delta_{i,f,c}, \quad \forall\, e \in E, f \in E, i \in I(e) \cap I(f), c \in C, \tag{10}$$

$$\sum_{d\in D, e\in d^P} \beta_{d,e,c} \leq U\pi_{e,c}, \quad \forall\, e \in E, c \in C, \tag{11}$$

$$\sum_{e\in E, i\in I(e), c\in C} \delta_{i,e,c} \leq l_i \gamma_i, \quad \forall\, i \in N, \tag{12}$$

$$\sum_{e\in E, i\in I(e), c\in C} \omega_{i,e,c} \leq l_i \gamma_i, \quad \forall\, i \in N. \tag{13}$$

Eqs. (9) and (10) state that whenever a termination occurs on a pair of incident edges, the corresponding termination variable δ must be 1. Analogously, Eq. (11) states that whenever a wavelength is used, the corresponding π must be set to 1. Eqs. (12) and (13) force the model to add a ROADM whenever a wavelength is either terminated or started (i.e, new circuits are added), respectively.

2.6. Signal Loss Recovering Constraints

Since a ROADM weakens signals and produces additional noise, the number of amplifiers and regenerators deployed in the system must be increased in presence of ROADMs.

Given path P, let $s(P) = \max_{e \in P}\{s_e\}$ be the maximum loss on path P, where s_e is the loss on edge e. Consider wavelength c and two consecutive terminations of this wavelength, i.e., between these two terminations, the content of the wavelength does not change. The signal loss along path P depends on the maximum loss $s(P)$ among all edges in P and the number k of ROADMs along P. Note that each ROADM on P degrades the signal even for wavelengths that are not terminated. On the other hand, whenever a wavelength is terminated, the signal can be regenerated and amplified with no extra cost.

Given maximum loss s, and number of ROADMs k on the path between two consecutive terminations of a wavelength, let $f(s, k)$ be the maximum number of edges between the two consecutive terminations of a wavelength, where $k = 0, 1, \ldots, K$ with K being the maximum number of ROADMs on any sup-path. We assume $f(s, k) \geq f(s, k+1)$ for any s and $k, 0 \leq k \leq K-1$, which expresses the fact that additional ROADMs further degrade the signal.

We model the signal loss concept by means of imposing a termination if the loss condition is not satisfied. Consider arbitrary path P and a possible wavelength c. If

(1) wavelength c is used on path P, and
(2) there are k ROADMs on path P, and
(3) $f(s(P), k) \leq |P|$, then

wavelength c must be terminated on P. If condition (3) is satisfied (recall that $f(s(P), k)$ is the maximum number of edges without requiring a termination), and wavelength c is not terminated on P, then this violates the definition of f. As a result, under these conditions a termination must occur.

For signal loss on wavelength c to be applicable on path P, all wavelengths must be used along path P. Let $|P|$ be the number of edges in P. Eq. (14) then captures condition (1).

$$\sum_{e \in P} \pi_{e,c} \leq |P| - 1 + \zeta_{P,c}, \quad \forall \text{ path} P, c \in C. \tag{14}$$

In (14), if $\pi_{e,c} = 1$ for every edge $e \in P$, then $\zeta_{P,c} = 1$. If at least one π on path P is 0, then $\zeta_{P,c}$ can be anything. These constraints imply consistency between $\pi's$ and $\zeta's$.

To consider condition (2), Eq. (15) is then included to capture the number of internal ROADMs along a path.

$$\sum_{i \in P^0} \gamma_i \geq k - k\theta_{P,k-1}, \forall \text{ path } P, f(s(P), k) \leq |P|, 1 \leq k \leq \min\{|P| - 1, K\}, \tag{15}$$

where P^0 is the set of all internal nodes in P.

In (15), if $\sum_{i \in P^0} \gamma_i = R$ for $R \leq |P| - 2, R \leq K$ (note that the two endpoints are excluded in P^0), then $\theta_{P,k}$ can be anything for $k \leq R - 1$. On the other hand, for $k \leq R$ we must have $\theta_{P,k} = 1$. This constraint is listed only for path P with $f(s(P), k) \leq |P|$ based on condition (3).

The actual loss requirement states that if wavelength c on path P satisfies $f(s(P), k) \leq |P|$ with k being the number of ROADMs on the internal nodes of P, then somewhere along path P wavelength c must be terminated (and thus regenerated) based on condition (3).

$$\sum_{i \in P^0, i \in I(e)} \delta_{i,e,c} \geq \theta_{P,k} + \zeta_{P,c} - 1,$$
$$\forall \text{ path } P, f(s(P), k) \leq |P|, c \in C, 0 \leq k \leq \min\{|P| - 2, K\}. \tag{16}$$

Eq. (16) is active only if $\zeta_{P,c} = \theta_{P,c,k} = 1$, which means that wavelength c is used on path P and it has k ROADMs internally. In this case, the right-hand side in (16) is 1, which implies that the wavelength must be terminated along path P.

The number of ROADMs along a wavelength must be bounded, and the number of edges a wavelength can use between two consecutive ROADMs must also be constrained.

$$\sum_{k=0}^{K} \theta_{P,k} \geq 1, \quad \text{for any path } P \text{ with } |P| - 2 \geq K, \tag{17}$$
$$\sum_{i \in P^0} \gamma_i \geq \zeta_{P,c}, \quad \text{for any path } P, |P| = M, c \in C. \tag{18}$$

Eq. (17) imposes that on a path used by the same wavelength, there must be at most K ROADMs. Eq. (18) imposes that on longer paths used by the same wavelength there must be at least 1 ROADM, with M as an upper bound constant on the number of edges in a wavelength between two ROADMs. Note that two ROADMs might not regenerate the signal.

Moreover, several constraints can be disaggregated. For example, constraint (11) can be written as:

$$\beta_{d,e,c} \leq U \pi_{e,c}, \quad \forall\, e \in E, \forall c \in C, d \in D, e \in d^P,$$

which provides a stronger LP relaxation, but it increases the size of the formulation. Other constraints that can be disaggregated are (8), (12), and (13). On the other hand, Eqs. (6), (7), (9), and (10) are presented in the disaggregated form.

Ideally all of the constraints should be disaggregated, but this could require excessive memory. The most appropriate formulation depends on the underlying computing architecture.

3. Three-Phase Algorithm

As noted in the previous section, even for mid-size instances the MIP model results in a considerably large-scale mathematical program. In terms of number of variables and constraints there is an exponential increase. For example, at least one constraint of type (14)–(18) may be required in the worst case for each path, thus easily resulting in a number ranging in millions. This strongly supports the application of the three-phase decomposition algorithm to effectively handle the underlying complex mathematical model.

The proposed heuristic decomposes the entire model into three tractable subproblems. The first subproblem, solved in Phase 1, is to determine the ROADM locations by only considering signal loss. In this phase, packing and blocking are neglected, and the model aims at determining all sites where a ROADM is required. As a result, all the corresponding regeneration locations for each service are selected. The problem does not decompose by service since the goal is to select regeneration sites for several services at the same node, which requires a single ROADM.

The second subproblem is solved in Phase 2. In this phase, all sites requiring ROADMs and regenerations are fixed for every service with model (1)–(13) being solved. Note that beyond Phase 1, signal loss requirements are always met. Since the resulting model is too large, we aggregate intermediate nodes to one 'artificial' node (see Figure 5). Intermediate nodes are nodes of degree 2 with no service requirements.

After solving the aggregated model and obtaining packing and blocking decisions on paths with wavelength sets on edges, the third subproblem is solved in Phase 3. In this phase, we disaggregate the wavelength decision variables and obtain the wavelength assignments for the whole network. A complete description of each phase is provided next.

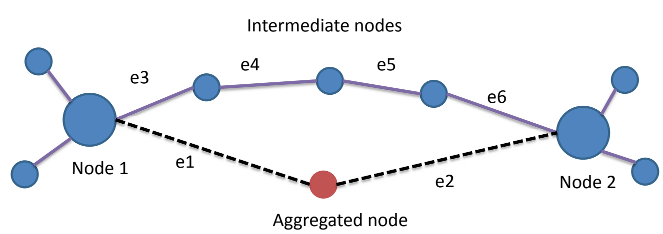

Figure 5: Node aggregation

3.1. Phase 1: ROADM Selection

The ROADM selection phase is to find the most cost efficient placement of line cards and ROADMs that satisfy signal loss requirements. Therefore, in this phase, we isolate the decision variables and constraints only related to signal loss, and build a smaller mixed integer model.

Several services come from the same client and they differ only with respect to the required bandwidth, i.e., v_d. As a result, all services from the same client use the same route d^p, which are subject to identical signal loss requirements. For this reason, we call all such services a service bundle. Formally, a service bundle is a set of all services with the same route.

Let $\bar{D}$ be the set of all service bundles. For each $\bar{d} \in \bar{D}$, there is a set $H_{\bar{d}} \subseteq D$ of services comprising the service bundle. All services in $H_{\bar{d}}$ have the same route, i.e., if $d_1 \in H_{\bar{d}}, d_2 \in H_{\bar{d}}$ for some service bundle $\bar{d}$, then $d_1^P = d_2^P$. We denote the set of all of the nodes in this unique path by $\bar{d}^P$. Sets $H_{\bar{d}}$ over all $\bar{d}$ partition D.

Given two nodes $i \in \bar{d}^P, j \in \bar{d}^P$, let $P(i,j,\bar{d}) \subseteq \bar{d}^P$ be the set of all nodes on path $\bar{d}^P$ between nodes i and j, including i and j. Under the assumption of this phase, i.e., no packing or blocking, constraints (14)–(18) do not need to be listed for all subpaths, but only for subpaths of $\bar{d}^p$. For each $\bar{d} \in \bar{D}$, we define the bandwidth of service bundle $\bar{d}$ as $b_{\bar{d}} = \sum_{d \in H_{\bar{d}}} v_d D_d$.

The model uses the same ROADM selection variables γ_i. For each service bundle $\bar{d} \in \bar{D}$ and node $i \in N$, we need to introduce new regeneration variables

$$\sigma_{i,\bar{d}} = \begin{cases} 1 & \text{if service bundle } \bar{d} \text{ is regenerated at node } i, \\ 0 & \text{otherwise.} \end{cases}$$

In order to capture the number of ROADMs on a segment of a route, for each $\bar{d} \in \bar{D}$, $i \in \bar{d}^P$, $j \in \bar{d}^P$, $0 \le k \le |\bar{d}^P| - 2$, we introduce variables

$$\kappa_{i,j,\bar{d},k} = \begin{cases} 1 & \text{if there are } k \text{ ROADMs on path } P(i,j,\bar{d}), \\ 0 & \text{otherwise.} \end{cases}$$

The objective function minimizes the weighted cost of all ROADMs and line cards:

$$\min \sum_{\substack{i \in N \\ \bar{d} \in \bar{D}}} b_{\bar{d}} CT_i \sigma_{i,\bar{d}} + \sum_{i \in N} CR_i \gamma_i, \tag{19}$$

where $b_{\bar{d}}$ provides a weight with respect to the bandwidth of a bundle.

The constraints are as follows.

- A ROADM is required at nodes with at least one line card. These constraints link together different bundles.

$$\sigma_{i,\bar{d}} \le \gamma_i, \quad \forall\, i \in N, \bar{d} \in \bar{D}. \tag{20}$$

- The number of ROADMs on a subpath must be captured. These constraints are similar to constraints (15) with variables θ replaced by κ.

$$\sum_{p \in P(i,j,\bar{d}) \setminus \{i,j\}} \gamma_p \ge k - k\kappa_{i,j,\bar{d},k-1}, \quad \forall\, \bar{d} \in \bar{D}, i \in \bar{d}^P, j \in \bar{d}^P,$$
$$f(s(P(i,j,\bar{d})), k) \le |P(i,j,\bar{d})|,$$
$$1 \le k \le |P(i,j,\bar{d})| - 1. \tag{21}$$

- Signal loss on service bundles is the main requirement of this phase. These constraints follow the reasoning of constraints (16) with variables θ replaced by κ. In addition, since we do not explicitly address wavelengths, variables ζ are no longer present.

$$\sum_{p \in P(i,j,\bar{d}) \setminus \{i,j\}} \sigma_{p,\bar{d}} \geq \kappa_{i,j,\bar{d},k}, \quad \forall\, \bar{d} \in \bar{D}, i \in \bar{d}^P, j \in \bar{d}^P, \tag{22}$$
$$f(s(P(i,j,\bar{d})), k) \leq |P(i,j,\bar{d})|,$$
$$0 \leq k \leq |P(i,j,\bar{d})| - 2.$$

- The maximum-number-of-ROADMs-on-a-subpath constraints are equivalent to (17).

$$\sum_{k=0}^{K} \kappa_{i,j,\bar{d},k} \geq 1, \quad \forall\, \bar{d} \in \bar{D}, i \in \bar{d}^P, j \in \bar{d}^P, |P(i,j,\bar{d})| - 2 > K. \tag{23}$$

- These maximum-number-of-segments-between-two-consecutive-ROADMs requirements replace (18).

$$\sum_{k \in P(i,j,\bar{d}) \setminus \{i,j\}} \gamma_k \geq 1, \quad \forall\, \bar{d} \in \bar{D}, i \in \bar{d}^P, j \in \bar{d}^P, |P(i,j,\bar{d})| = M. \tag{24}$$

A desirable feature of this model is the fact that the number of constraints is limited, i.e., it is polynomial in the input size.

3.2. Phase 2: Aggregation

By solving the phase 1 model, we obtain optimal variables γ^* and σ^*. In the second phase we fix the ROADM variables γ to the values γ^* for those nodes with $\gamma_i^* = 1$. All of the steps that follow impose that terminations happen at locations i with $\sigma_{i,\bar{d}}^* = 1$ for each service d with $d \in H_{\bar{d}}$. This guarantees signal loss requirements for all subsequent steps.

Figure 5 shows two nodes, labeled by node 1 and node 2, connected by a path with 3 nodes, each with degree 2, and with no service requirements. Intuitively, it seems that on this path all packings and blockings can be done at a single node. We actually provide later an analytical statement that there is no loss of optimality by collapsing the three nodes into a single one assuming that packing and blocking operations are decoupled. Since packing and blocking can occur concurrently, this aggregation is not without loss of optimality.

For the aggregated model, a new graph $\bar{G}(\bar{N}, \bar{E})$ is introduced. Set $\bar{N}$ is the set of all aggregated nodes, nodes with add/drop service requirements, and nodes with degree larger than 2. Set $\bar{E}$ indicates the set of edges. All consecutive nodes with degree 2 are collapsed into a single node. We use similar decision variables, constraints, and the objective function except that they are defined with respect to $\bar{G}$.

For each aggregated node $i \in \bar{N}$, let $S(i) \subseteq N$ be the set of original nodes aggregated to i. Given γ^* and σ^* from Phase 1, which are based on the original graph, we define $\gamma^*_i = 1$, $\sigma^*_{i,\bar{d}} = 1$ for an aggregated node $i \in \bar{N}$ if there exists $j \in S(i)$ with $\gamma^*_j = 1$, $\sigma^*_{j,\bar{d}} = 1$, respectively.

To capture signal loss correctly, we solve the model described by constraints (1)–(10), (12), (13) with the following additions. For each $c \in C, d \in D$, and edge $e \in d^P$, we define new decision variables

$$\vartheta_{d,e,c} = \begin{cases} 1 & \text{if wavelength } c \in C \text{ on edge } e \text{ includes a} \\ & \quad \text{circuit from service } d \\ 0 & \text{otherwise.} \end{cases}$$

To properly define these variables we add the following constraints:

$$\beta_{d,e,c} \leq m_d \vartheta_{d,e,c} \qquad d \in D, e \in d^P, c \in C.$$

In addition, for each $i, \bar{d} \in \bar{D}$ with $\sigma^*_{i,\bar{d}} = 1$, and each $d \in H_{\bar{d}}$ we impose

$$\vartheta_{d,c,e} \leq \delta_{i,e,c} \qquad c \in C, i \in I(e), e \in d^P.$$

These constraints guarantee signal loss requirements since they specify that any wavelength including a service is terminated if the solution from Phase 1 specifies so. The model will use these facts to possibly repack or block wavelengths at such pre-determined terminations.

It is expected that network $\bar{G}$ is much smaller than the original network G and thus this new model is better manageable.

3.3. Phase 3: Disaggregation

After solving the subproblem in Phase 2, we need to disaggregate the solution to obtain a complete solution. This can again be done by using Phase 2 for each aggregated node. For each aggregated node $k \in \bar{N}$, we consider the underlying path defined by $S(k)$ in isolation. In addition, wavelength compositions on the first and last edge on the path are fixed based on the solution from Phase 2. Consider the top configuration in Figure 6. The solution on the lower path is obtained by Phase 2 and it yields an optimal wavelength assignment sets c_1 on edge e_1, and c_2 on edge e_2. By fixing the wavelength assignment sets c_1 and c_2 on edges e_3 and e_6, we can determine the wavelength sets on each edge on the path in the disaggregated model. The model to carry this out is identical to the model in Phase 2 except that it is applied to the path induced by $S(k)$. If this is executed for each aggregated node, we obtain a feasible solution to the original network.

We devote the rest of this section to state two propositions that justify our aggregation technique. In the statements, we assume that signal loss requirements are not needed, i.e., no σ^* is used. Let us fix an aggregated node.

Proposition 1. *If only blocking is allowed, then the cost of the disaggregated solution is at least as good as the cost of the aggregated solution.*

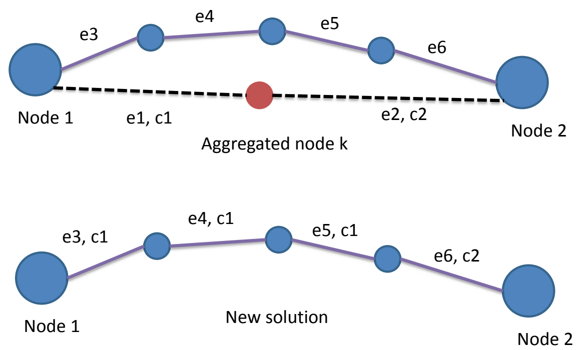

Figure 6: Blocking in the disaggregated network

Proof. As shown in Figure 6, when we fix the wavelength set of the first edge to c_1, and the wavelength set of the last edge to c_2, and set the wavelength sets of intermediate edges to c_1, we obtain a solution to the disaggregated model with no higher cost. □

Proposition 2. *If only packing is considered, then the cost of the disaggregated solution is at least as good as the cost of the aggregated solution.*

Proof. We prove this by induction on the number n of intermediate nodes in $S(k)$. The base case $n = 2$ is actually the bulk of the work.

Assume that we have only two intermediate nodes i and j, as shown in Figure 7. We show this case by induction on the number of circuits involved on edge e_1 (this is a nested induction within the overarching induction on the number of nodes). Note that this number is the same on the remaining edges e, f, g and f_1.

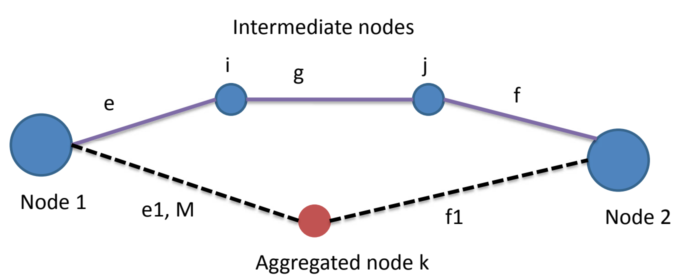

Figure 7: Packing in the disaggregated network

If there is a single circuit, the claim is trivial since in the disaggregated and aggregated models there are no terminations. Let us assume now that we have a set M of circuits in question.

Given a set M of circuits, let us denote by $u(M)$, $v(M)$ the number of line cards in the aggregated and disaggregated models, respectively. We need to show that $u(M) \leq v(M)$.

Let us arbitrarily select one of the circuits in M, which for reference is labeled as q and it belongs to wavelength $t_1 \in M$. Next we remove circuit q

from consideration and we apply the induction step, which yields $u(M\backslash q) \leq v(M\backslash q)$.

Next we study possible cases the way circuit q is inserted back into consideration. It is not difficult to see that an addition of a single new circuit to the aggregated model implies $u(M) \in \{u(M\backslash q), u(M\backslash q) + 2, u(M\backslash q) + 4\}$. If $u(M) = u(M\backslash q)$, then the claim follows since $u(M\backslash q) \leq v(M\backslash q) \leq v(M)$. The addition of new circuits can only increase the number of line cards in both models.

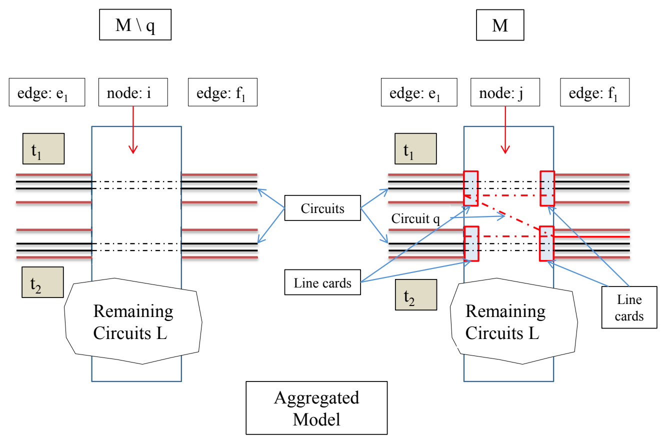

Figure 8: Case $u(M) = u(M\backslash q) + 4$ in the aggregated model

Case $u(M) = u(M\backslash q) + 4$ is depicted in Figure 8. It is not difficult to observe that the case depicted is the only case for this to happen. By induction we have $u(L) \leq v(L)$ where L are all circuits disjoint from those in wavelengths t_1 and t_2. The remaining two wavelengths under the aggregated model in the case of $M\backslash q$ do not require a line card, while there is a need for 4 line cards in the case of M. Figure 9 shows a possible case under the disaggregated model. We note that the number of line cards under M is increased by 4. This observation implies the statement in this case. The remaining two cases shown in Figure 10 can be argued in a similar way.

Case $u(M) = u(M\backslash q) + 2$ is illustrated in Figure 11. It can be treated in the same fashion as the previous case. We conclude that the claim holds for two nodes.

To complete the overall induction, we assume Proposition 2 holds when the number of intermediate nodes is less than or equal to l, $l \geq 3$. When $n = l + 1$, we can select two adjacent nodes, and aggregate them into one node. As shown in Figure 12, the two adjacent nodes $node_1$ and $node_2$ are aggregated into $node_3$. Consequently, the new setting containing $node_3$ has l intermediate nodes. Based on the proof of the base case $n = 2$, the number of line cards needed at $node_1$ and $node_2$ is larger than or equal to the number of line cards needed at $node_3$. Therefore, the total number of line cards needed for $n = l + 1$ is larger than or equal to the total number of line cards needed when $n = l$.

This completes the overall proof. $\square$

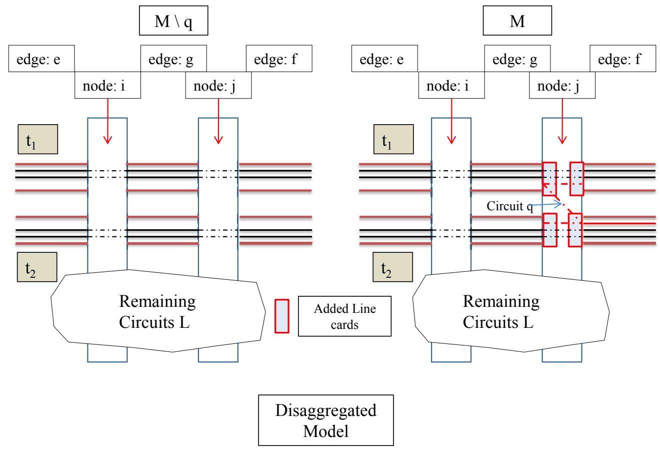

Figure 9: Case $u(M) = u(M\backslash q) + 4$ in the disaggregated model

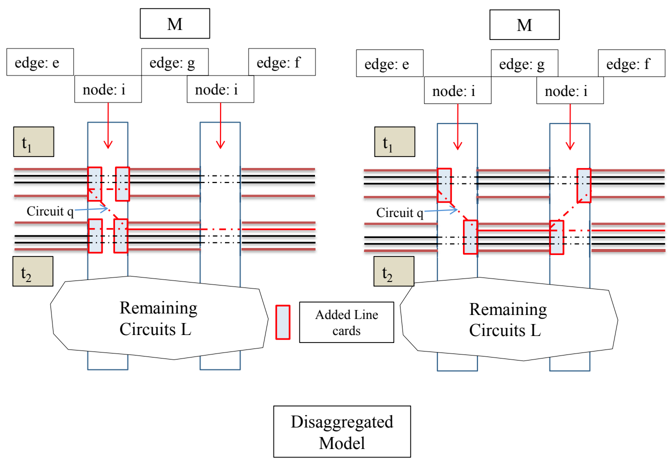

Figure 10: All other possible cases $u(M) = u(M\backslash q) + 4$ in the disaggregated model

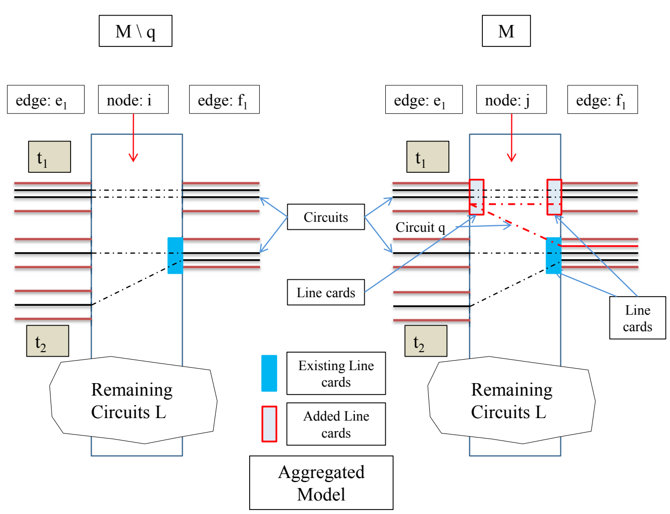

Figure 11: Case $u(M) = u(M\backslash q) + 2$ in the aggregated model

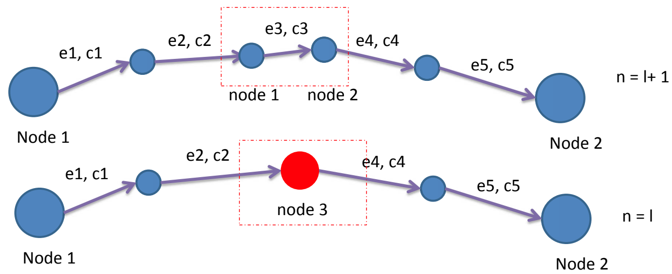

Figure 12: Induction: Aggregated nodes when $n = l + 1$.

4. Computational Experiments

In this section we present and discuss the results of running the three-phase algorithm described in Section 3. The experiments aim at providing a complete analysis of the performance of the algorithm.

All experiments reported were conducted on the set of test instances shown in Table 1, where an identifier for each test instance is given in Column 1. Columns 2-4 show the networks' properties in terms of services, nodes and edges, respectively. Furthermore, the number of wavelengths is set to 40 for each test case in Table 1. Instances A-E were constructed manually and their traffic scenarios were generated randomly. Instances F and G, on the other hand, correspond to a network topology provided by an optic cable provider with presence in the entire U.S. and it reflects their true topology and cost parameters. The provider also uses 40 wavelengths. All engineering requirements such as loss were provided by the company. To protect our confidentiality agreement, actual data and exact numerical results are provided for all but instances F and G. For these two real-world instances, we provide range values and conduct further numerical analyses.

Table 1: Test instances

	Number of		
Ref	Services	Nodes	Edges
A	20	122	159
B	15	112	149
C	50	126	163
D	10	305	327
E	15	417	450
F	[38,43]	493	536
G	[72,80]	493	536

We conducted the experiments on a 3.2GHz Intel(R) Pentium IV dual-core central processing unit with 8 GByte RAM. The algorithm was implemented in Python and the mathematical models were solved by calling MIP optimizer Gurobi 2.0 [14]. We impose a time limit of 3600 CPU-seconds on each instance.

4.1. Overall Performance

In addition to (19), which weighs the cost of a line card with bandwidth, we present the following alternative specifications to the model of Phase 1:

$$\min \sum_{\substack{i \in N \\ \bar{d} \in \bar{D}}} CT_i \sigma_{i,\bar{d}} + \sum_{i \in N} CR_i \gamma_i, \tag{25}$$

$$\min \sum_{\substack{i \in N \\ \bar{d} \in \bar{D}}} b_{\bar{d}} CT_i \sigma_{i,\bar{d}} + \sum_{i \in N} s_i CR_i \gamma_i, \tag{26}$$

where $s_i = \sum_{\substack{i \in \bar{d}^P \\ \bar{d} \in \bar{D}}} b_{\bar{d}}$.

We call (25), (26), (19) cost-, bandwidth-, and cost-bandwidth-based objective functions, respectively.

Our main goal in applying three different objective functions is to evaluate the quality of the approximate solutions provided by Phase-1, and thus assessing the overall performance of the three-phase algorithm. Hence, (25), (26), (19) basically provide Phase-2 and Phase-3 with different weights to the cost of ROADMs and line cards.

Tables 2–4 show the results of applying the three objectives to the set of test instances shown in Table 1. For each instance, the required number of ROADMs, line cards, and objective values (in millions) are shown in Columns 2–4, respectively.

Table 2: Results based on the cost-bandwidth-based objective

Ref	No. of ROADMs	No. of Line Cards	Obj ($\times 10^6$)
A	55	240	3.10
B	65	328	3.39
C	71	394	7.52
D	60	118	1.81
E	79	153	2.70
F	[60,75]	[520,600]	[6.1,9.4]
G	[90,125]	[820,890]	[8.0,11.0]

Table 3: Results based on the cost-based objective

Ref	No. of ROADMs	No. of Line Cards	Obj ($\times 10^6$)
A	58	246	3.45
B	67	345	3.59
C	75	402	7.45
D	64	121	1.85
E	79	140	2.76
F	[65,85]	[530,615]	[6.5,9.5]
G	[97,125]	[780,870]	[8.5,11.5]

Table 4: Results based on the bandwidth-based objective

Ref	No. of ROADMs	No. of Line Cards	Obj ($\times 10^6$)
A	58	255	3.26
B	72	339	3.60
C	77	410	7.82
D	70	125	1.88
E	86	160	2.76
F	[70,95]	[535,620]	[6.7,9.7]
G	[110,135]	[840,900]	[7.5,10.5]

We use cost-bandwidth-based objective as a benchmark and compare it against the other two in terms of the number of ROADMs, line cards, and objective values provided for instances A–E.

Relative improvement values are given in Table 5. When comparing cost-bandwidth vs. cost, Table 5 shows reductions up to 6.7%, 5.2% and 11.3% of cost-bandwidth in the number of ROADMs, line cards, and the overall cost, respectively. Regarding bandwidth solutions, reductions up to 16.7%, 6.3% and 6.2% of cost-bandwidth in the same areas, respectively, are observed.

Table 5: Relative improvements on non-real-world instances

	cost-bandwidth solutions vs.					
	cost solutions			bandwidth solutions		
Ref	ROADMs	Line Cards	Obj	ROADMs	Line Cards	Obj
A	5.5%	2.5%	11.3%	5.5%	6.3%	5.2%
B	3.1%	5.2%	5.9%	10.8%	3.4%	6.2%
C	5.6%	2.0%	-0.9%	8.5%	4.1%	4.0%
D	6.7%	2.5%	2.2%	16.7%	5.9%	3.9%
E	0.0%	-8.5%	2.2%	8.9%	4.6%	2.2%

From the values in Table 5, we also observe that cost-bandwidth solutions improve cost solutions by an average of 4.2%, 0.8% and 4.1% with respect to the number of ROADMs, line cards, and the overall cost. Analogously, reductions of 10.0%, 4.8% and 4.3% on average are obtained within the same areas, respectively, when compared with bandwidth solutions. We clearly proclaim the cost-bandwidth-based objective to be the winner.

Table 6 shows the overall performance with respect to the computational time of the three-phase algorithm when applying the 3 different objectives to each instance. The performance is measured in terms of both the CPU time (in seconds) consumed in each phase and the resulting optimality gap of the underlying mathematical model. Note that this is not the optimality gap with respect to the optimal solutions of the entire problem but of each individual phase. Whenever an optimal solution is not provided within the time limit, the table shows the corresponding MIP gap in parentheses instead.

From Table 6 we observe that Phases 1 and 3 require a considerably lower computational time than Phase 2. For example, all but instances F and G require no more than 10 seconds and 5 minutes to converge in Phases

Table 6: Run times and MIP gaps of each phase of the proposed algorithm

Ref	Phase	CPU Time (sec) cost-bandwidth	cost	bandwidth
A	1	1	2	1
	2	(14.6%)	(13.1%)	(14.4%)
	3	19	5	66
B	1	0	0	1
	2	(7.1%)	(7.5%)	(7.5%)
	3	64	1	63
C	1	2	2	2
	2	(45.1%)	(43.8%)	(44.2%)
	3	247	248	245
D	1	8	0	4
	2	(7.5%)	(7.5%)	(8.6%)
	3	69	68	68
E	1	1	0	0
	2	(9.6%)	(11.5%)	(9.8%)
	3	108	67	97
F	1	96	64	67
	2	(59.2%)	(57.4%)	(58.9%)
	3	1420	1302	1298
G	1	540	336	422
	2	(55.4%)	(62.2%)	(53.7%)
	3	(20.7%)	3182	2825

1 and 3, respectively, whereas for Phase 2 they reach the time limit before converging. The different objectives tested in Phase 1 proved to have an insignificant influence on this pattern. Yet, the cost-based objective slightly improves the run times in Phase 3 and reduces the MIP gaps in Phase 2 in comparison with the values observed from the remaining two objectives.

4.2. Real-World Instances

4.2.1. Bandwidth and Wavelength Utilizations

We investigate resource utilization in terms of bandwidth utilization and wavelength utilization with respect to the different objective functions tested on instances F and G. Bandwidth utilization is defined as

$$\text{Bandwidth Utilization} = \frac{\sum_{e\in E}\sum_{c\in C}\sum_{d\in D} v_d \beta_{d,e,c}}{\sum_{e\in E}\sum_{c\in C} U} \cdot 100\%,$$

while wavelength utilization on an edge is defined as

$$\text{Wavelength Utilization} = \frac{40 - \text{number of unused wavelengths}}{40} \cdot 100\%.$$

Table 7 shows the bandwidth utilization and the average wavelength utilization, including the standard deviation for wavelength utilization. Note that in our case the maximum number of wavelengths is set to 40. We can observe from the table that instance G yields a more balanced wavelength utilization, i.e., its standard deviation is around 0.076. Since the utilization should be high, this table confirms that cost-bandwidth is the most effective option.

Table 7: Bandwidth and wavelength utilizations

		Effective use of		Standard
Ref	Objective	Bandwidth (%)	Wavelength (%)	Deviation
F	Cost-based	85%	83%	0.158
G		74%	86%	0.076
F	Bandwidth	85%	83%	0.157
G		73%	86%	0.078
F	Cost-Bandwidth	87%	84%	0.156
G		74%	86%	0.076

4.2.2. Terminations at Nodes

With the cost-bandwidth-based objective function, in Figure 13 we show histograms of the number of terminations (packing and blocking) for instances F and G.

For instance F, the model tends to place more line cards at individual nodes, i.e., most nodes have more than or equal to 10 terminations. This is also reasonable for the number of services required by this instance. For instance G, we observe that most of the nodes have less than or equal to 10 terminations. This is expected since a large network offers more choices to place line cards, thus it tends to spread line cards more evenly. On average our proposed solution requires 9 to 11 line cards per node.

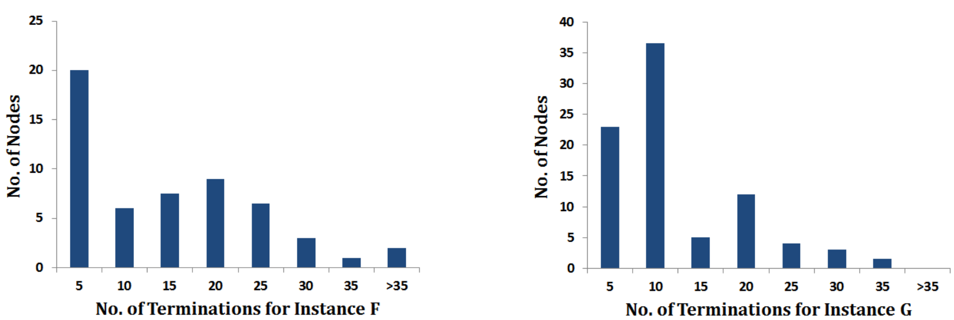

Figure 13: Histograms of terminations for the bandwidth-based objective function

5. Conclusions

In this paper, we have proposed a heuristic approach to effectively tackle the optimal placement problem of ROADMs and line cards when designing fiber optic networks. More precisely, we introduced a novel mixed integer programming model as the main core of the proposed approach, where the signal loss, wavelength packing and blocking are captured. Since the underlying model is computationally challenging, we designed and implemented a three-phase algorithm to tackle it. The algorithm efficiently provides good feasible solutions by dividing the entire problem into three subproblems, which are relatively easy to solve. We provide theoretical justifications for the key idea behind this approach. To assess the performance of the algorithm, we conducted computational experiments on a set of test instances of different sizes. We also present different objective functions to a model in a phase that yields different overall solutions.

The contributions of this research are fourfold. First, we solve a relevant industrial problem for a very competitive industry. Second, to the best of our knowledge, the literature reveals no research work on tackling signal degradation and loss when solving the ROADM placement problem. Third, we incorporate the termination (blocking and packing) into our model. Due to its inherent complexity, the research community has overlooked terminations resulting from the nature of ROADMs and line cards. Finally, we have implemented a three-phase algorithm for this problem, which proved to be efficient on a set of test instances, including two real-world cases. The algorithm has been tested on real-world instances and exact engineering rules provided by a telecommunication company.

References

[1] K. Zhu, B. Mukherjee, A review of traffic grooming in WDM optical networks: Architectures and challenges, Optical Network Magazine (2003) 55–64.

[2] A. Sutter, F. Vanderveck, L. Wolsey, Optimal placement of add/drop multiplexers: Heuristic and exact algorithms, Operations Research, Vol. 46 (1998) 719–728.

[3] G. Belvaux, N. Boissin, A. Sutter, L. A. Wolsey, Optimal placement of add/drop multiplexers: Static and dynamic models, European Journal of Operational Research, Vol. 108 (1998) 26–35.

[4] C. Lee, S. Koh, Assignment of add-drop multiplexer (ADM) rings and digital cross-connect system (DCS) mesh in telecommunication networks, Journal of the Operational Research Society, Vol. 52 (2001) 440–448.

[5] B. Fortz, P. Soriano, C. Wynants, A tabu search algorithm for self-healing ring network design, European Journal of Operational Research, Vol. 151 (2003) 280–295.

[6] O. Goldschmidt, A. Laugier, E. V. Olinick, SONET/SDH ring assignment with capacity constraints, Discrete Applied Mathematics, Vol. 129 (2003) 99–128.

[7] E. Macambira, N. Maculan, C. de Souza, A column generation approach for SONET ring assignment, Networks, Vol. 47 (2006) 157–171.

[8] G. Calinescu, P. Wan, Splittable traffic partition in WDM/SONET rings to minimize SONET ADMs, Theoretical Computer Science, Vol. 276 (2002) 33–50.

[9] O. Gerstel, P. Lin, G. Sasaki, Wavelength assignment in a WDM ring to minimize cost of embedded SONET rings, INFOCOM, Vol. 1 (1998) 69–77.

[10] H. Holler, S. Voβ, A heuristic approach for combined equipment-planning and routing in multi-layer SDH/WDM networks, European Journal of Operational Research, Vol. 171 (2006) 787–796.

[11] Z. Hongyue, M. Biswanath, Online connection provisioning in metro optical WDM networks using reconfigurable OADMs, Journal of Lightwave Technology, Vol. 23 (2005) 2893–2901.

[12] J. Fourneau, N. Izri, D. Verchère, Analysis and optimization of aggregation in a reconfigurable optical add/drop multiplexer, Lecture Notes in Computer Science, Vol. 5764 (2009) 120–131.

[13] H. Wei, B. Ningfeng, S. Xiaohan, Design of flexible and reconfigurable optical add/drop multiplexer, Journal of Optical Fiber Technology, Vol. 17 (2011) 247–251.

[14] `Gurobi Optimization Inc.`, `http://www.gurobi.com/` (2010).