Resource-aware provisioning strategies in translucent elastic optical networks

Original
Resource-aware provisioning strategies in translucent elastic optical networks / Gonzalez-Montoro, N; Finochietto, Jm; Bianco, A. - In: COMPUTER COMMUNICATIONS. - ISSN 0140-3664. - 180:(2021), pp. 134-145. [10.1016/j.comcom.2021.09.008]

Publisher:
ELSEVIER

Published
DOI:10.1016/j.comcom.2021.09.008

Resource-Aware Provisioning Strategies in Translucent Elastic Optical Networks

Nehuen Gonzalez-Montoro[a], Jorge M. Finochietto[a,b,d], Andrea Bianco[c]

[a] *Universidad Nacional de Córdoba, Facultad de Ciencias Exactas, Físicas y Naturales, Laboratorio de Comunicaciones Digitales. Córdoba, Argentina*
[b] *CONICET, Instituto de Estudios Avanzados en Ingeniería y Tecnología (IDIT). Córdoba, Argentina*
[c] *Politecnico di Torino, Dipartimento di Elettronica e Telecomunicazioni. Torino, Italy*
[d] *Corresponding author: jorge.finochietto@unc.edu.ar, Laboratorio de Comunicaciones Digitales - FCEFyN - UNC - Av. Vélez Sarsfield 1611 - Córdoba (5000) - Argentina*

Abstract

Elastic optical networks (EONs) enable a more efficient use of spectrum resources than traditional wavelength-division multiplexing (WDM) networks, mainly by means of a flexible frequency grid and the use of bandwidth variable transponders (BVTs). In WDM networks, regeneration is avoided because it always produces a cost increase and it is only recommended when the transmission length cannot be met. Instead, in EONs, a trade-off exists between regeneration and spectrum use. Indeed, blocking probability can be widely reduced by incorporating regeneration (BVTs in a back-to-back configuration) to shorten the transmission lengths, therefore increasing spectrum efficiency. Provisioning algorithms that take into account this trade-off consider regeneration either as a cost to be minimized (full regeneration capacity) or as a resource to be smartly used (bounded number of predeployed transponders).

In recent years, algorithms that minimze the resource use of either spectrum or BVT have been proposed. These algorithms, however, penalize the overall network performance in terms of blocking probability and cost, since they do not take into account the available resources. In this work, we propose two novel *resource-aware* algorithms, that take into account all available resources in the provisioning process. We report simulation results showing that using these *resource-aware* algorithms: i) when full regeneration capacity is assumed, the regeneration cost can be reduced by more than a 30% compared to the opaque solution and by a 10% compared to the best known algorithms; ii) for the bounded case, resource-aware strategies can reduce blocking probability by thousands of times compared to the transparent solution and by hundreds of times compared to the best known algorithms. Our results show that resource awareness helps to decrease the blocking probability while increasing the actual network capacity, thanks to a more efficient assignment of both spectrum and BVT resources.

Keywords: Elastic Optical Networks, Provisioning, Translucent, Regeneration, Resource Aware

1. Introduction

Wavelength division multiplexing (WDM) and dense wavelength division multiplexing (DWDM) networks base their operation on a fixed 100/50 GHz grid [1]. Even if these networks can transport large bandwidths they become inefficient in presence of heterogeneous and variable traffic demands. The elastic optical network (EON) architecture [2, 3], proposed almost a decade ago and extensively studied since then [4], enables a more efficient use of spectrum resources by means of a flexible grid structure. In this architecture, demands are provisioned using multiple adjacent 12.5GHz frequency slots. The amount of slots required by a single demand depends on the demanded bitrate, the transmission length, the modulation format, the baud rate, the number of carriers and the error correction overhead. Bandwidth-variable transponders (BVTs) can adjust the baud rate, the modulation format, the number of carriers and the error correction overhead to transport a demand with a specified bitrate minimizing the required frequency slots. This flexibility enables EONs to establish optical connections with “arbitrary” bandwidths (frequency slots) [5, 4].

On the other hand, regeneration has been widely used in optical networks to reduce the blocking probability when provisioning traffic demands. In traditional WDM networks, regeneration has been mainly proposed to overcome physical impairments that limit the maximum transmission distance [6]. Besides, regeneration can also be used to reduce blocking by providing wavelength conversion capabilities that can tackle spectrum fragmentation issues. However, regeneration in traditional WDM networks has no impact on the spectrum utilization because each lightpath exploits either a 50GHz or a 100GHz channel regardless of the path length or assigned wavelength. Instead, regeneration in EONs can also be used to compress the amount of spectrum required for provisioning a demand. To this end, a flexible regenerator composed of back-to-back bandwidth-variable transponders [7, 8, 9] can be used, where each transponder handles the same rate but not necessarily requires the same amount of frequency slots. These regenerators can be configured to support, independently on each carrier, a specific modulation format and baud-rate. The shorter the distance between regeneration points, the higher the modulation format that can be used, leading to an increase in spectral efficiency and spectrum saving. Flexible regeneration enables a trade-off between spectrum and transponder costs, as described in [10, 11].

The use of regeneration and its impact on blocking ratio has been widely studied in the past. Nevertheless, recently it has gained relevance mainly due to the trade-off between regeneration and spectrum use in EONs that we described in the previous paragraph. In recent years, the effort have focused on obtaining provisioning algorithms that take advantage of this trade-off [8, 12, 13, 14, 15, 16]. In this context, regenerators can be considered either as a cost to be minimized or as a resource to be smartly used. In the first case, spectrum is considered as the limiting resource and the use of regeneration to reduce the blocking probability translates to node cost increase. In the second scenario, regenerators are *bounded* to the amount of transponders that each node is equipped with, at network design phase. The goal becomes to efficiently use available transponders. If spectrum slot lack becomes the main bottleneck when accepting traffic demands, transponders can be used to save spectrum to reduce the demand blocking probability. In operational networks, regeneration capabilities can be increased much faster and smoothly by equipping nodes with more transponders, while increasing spectrum resources is typically slower and may require a huge investment due to fiber cable deployments or renting costs.

In this paper we evaluate different provisioning strategies that tackle both unbounded and bounded regeneration scenarios to illustrate the value of regeneration in EONs. We name these strategies as *resource-aware* strategies, because they aim at selecting the best combination of available spectrum and transponder resources for establishing a translucent lightpath. Based on our previous contribution [14], originally designed for the full regeneration capacity scenario, we propose a novel strategy, suited to both scenarios. The strategy differs from others proposed in the literature, which are typically agnostic to available resources. We demonstrate that resource awareness helps in decreasing the blocking probability and increases the actual network capacity. thanks to a more efficient assignment of both spectrum and transponder resources. The paper is organized as follows: in Sec.2 previous research activities are summarized and compared. In Sec.3 we analyze the problem of selecting an optimal provisioning candidate in Elastic Optical Networks. We present the used system model in Sec.4. Then we introduce the proposed resource-aware algorithms in Sec.5. Finally, in Sec.6 we show simulation results and in Sec.7 we derive the conclusions.

2. Related Work

Regeneration in optical networks has been widely studied in terms of Regeneration Placement (RP), where regeneration is mainly used to compensate for transmission length and/or spectrum fragmentation issues that preclude lightpath establishment on a given path. [17, 18]. Thus, the use of regeneration focus on reach extension and/or wavelength conversion as a mean to reduce the lightpath blocking probability. This problem has also been studied for EONs in [19, 20, 21] focusing on selecting in advance which nodes should support regeneration to overcome physical impairments. In [11] the impairment-aware translucent routing problem was studied for EONs. Flexible regeneration was modeled considering different modulation formats supported by regenerators, but their placement is chosen before the provisioning process starts. Results showed that regeneration can improve the blocking performance. In [10] the trade-off between regeneration placement and spectrum usage in EONs was studied for the first time. An offline spectrum assignment and regeneration placement heuristic was proposed. Results showed that there is a non-linear relationship between the overall spectrum usage and the overall regeneration use. Moreover, the authors showed a considerable reduction of spectrum usage on EON when regenerators are available.

An adaptive modulation and routing and spectrum allocation algorithm was proposed in [22] for regeneration-enabled EONs. The problem was split in two: the route, modulation and spectrum assignment problem (RMSA) and the regenerator placement problem (RP). The RMSA was addressed with an adaptive routing algorithm considering link utilization for routing, modulation penalization for modulation selection and regenerator penalization for regeneration selection. Results showed that the best trade-off between cost and blocking performance is obtained when adaptive modulation and regeneration selection are implemented. In [9] a MILP formulation of the RMSA problem on regeneration-enabled EONs was proposed. As in [11, 10, 23, 22], the regenerator placement problem was modeled as a separate problem. Results showed that there is an overall trade-off between the number of regenerators and the spectrum utilization.

In [7] the performance improvements obtained by regeneration and modulation conversion in translucent EONs was studied. Authors proposed a regenerator model that make use of additional transponders to increase the number of carriers in transparent sections of the translucent paths. They showed that a translucent path with more regeneration points can use less transponders than a less regenerated option. This happens due to the spectrum compression that high order modulation formats can obtain. An optimization approach was used and results showed that gains can be achieved in terms of served traffic by allowing regeneration with modulation conversion. Recently, the joint route, spectrum, modulation and regenerator assignment (RSMRA) problem has been studied for non-protected provisioning [24, 12, 25, 8, 13, 14, 15] and for protected provisioning [26] in EONS. In [27] RSMRA problem was studied for EONs with sliceable bandwidth variable Transponders and in [16] for spectral-spatially flexible optical networks.

In [14] we proposed an algorithm that aims at finding the optimal solution for the problem of minimizing the regeneration cost (expressed as the number of required transponders) when full regeneration capacity is assumed (i.e., nodes have enough transponders to provide regeneration at each intermediate node). However, since only minimizing the regeneration cost can waste spectrum resources, which can become the main blocking source, the algorithm assigns a spectrum budget (threshold) to each demand to avoid spectrum waste when regeneration can save significant spectrum resources. This algorithm has also been analyzed recently by other authors in [15], and its complexity has been compared to different existing heuristics.

In this paper we discuss how to overcome complexity limitations of our original algorithm, and propose a new strategy that offers better performance with similar complexity. This new contribution is a generalization of our previous algorithm, but it can also be used in bounded regeneration scenarios, where both resources, transponders and spectrum, are considered scarce.

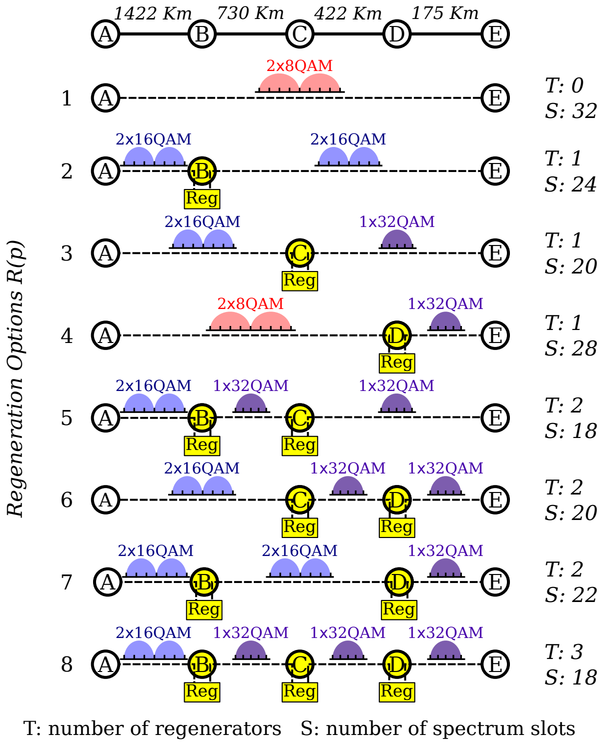

Figure 1: All possible provisioning options for a four hops path.

3. Optimal Provisioning Candidates

Multiple candidate solutions exist for provisioning a traffic demand over a path p. These solutions range from the transparent case (no regeneration in any node) to the opaque one (regeneration at each intermediate node), including all translucent solutions which implement regeneration at some intermediate nodes.

Each candidate solution can have different requirements in terms of network resources. The two main resources of an EON are spectrum and transponders. In this work, we focus on the optimization of these two resources. We assume that other resources as bandwidth variable spectrum switches are dimensioned in a way that they are not a limitation. It is straightforward that some solutions are better than others in terms of either spectrum or regeneration cost. Obviously, if two solutions have the same spectrum cost but one requires less regenerators, the last one should be preferred. However, solutions may be difficult to compare in general. For example, if one solution requires less spectrum but more regenerators than another one, it could argued which one is better. Indeed, both may be optimal solutions, depending on the optimality criteria. We defer our detailed discussion on how to select a solution from a set of optimal ones to Sec. 5. Instead, we first discuss how optimal solutions can be found and provide some initial insight on the selection criteria.

We define $R(p)$ as the set of all potential solutions for provisioning a traffic demand over path p. However, due to lack of spectrum and/or spectrum fragmentation or lack of transponders at intermediate nodes, not all solutions may be available in practice. We denote by $R'(p)$ the set of feasible candidate solutions. Each feasible solution $r \in R'(p)$ has a spectrum cost $S(r)$ and transponder cost $T(r)$ associated with, where $S(r)$ is the total number of frequency slots (computed as the sum of the required slots at each hop) required to provision a demand over path p using solution r , while $T(r)$ is the number of additional transponders used by the provisioning option r.

Given a path p, a set of Pareto optimal (or efficient) solutions $P(p) \subset R'(p)$ can be found for each new demand to be provisioned. This set ranges from solutions with low regeneration cost but with high spectrum utilization to others which demand more regeneration but require less spectrum resources. This set can be found solving the following multi-objective optimization problem:

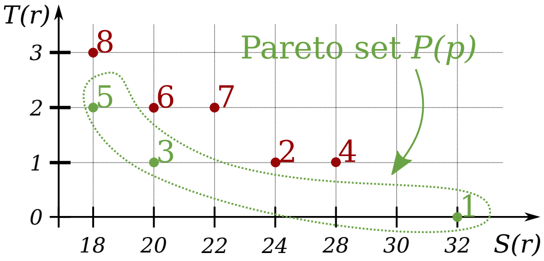

Figure 2: Provisioning options in the Pareto set

$$\begin{aligned} & min(T(r), S(r)) \\ & \quad \text{s.t. } r \in R(p) \end{aligned}$$

To illustrate this problem, we consider the example shown in Fig. 1, where a 400 *Gb/s* demand from node A to node E needs to be provisioned over a 5-node linear path (A, B, C, D, E) of total length 2750 *km*, using a 50 *GBaud* transponder with 25% FEC. All provisioning alternatives are reported: the transparent option (1), the opaque option (8) and all the translucent alternatives (2-7). Details on the transmission model used to derive these results are presented later in Sec.4.3. The transparent option (1) requires 2 carriers with a low order modulation format (8QAM) to achieve the full transmission length of the path. A total of $S(r) = 32$ slots are required, 4 slots per carrier at each of the 4 links. In contrast, the opaque solution (8) makes use of higher modulation formats in all hops, thus requiring much less frequency slots, $S(r) = 18$, at the cost of introducing 3 regenerators, $T(r) = 3$. Translucent options (2-7) require intermediate values of regenerators and frequency slots.

In Fig. 2, we plot all solutions of $R(p)$ in terms of their costs spectrum and transponder costs. All green solutions belong to the Pareto set $P(p)$, while those in red are not Pareto optimal, because are dominated by other solutions (i.e., there exist other solutions that perform better both in terms of slots and number of regenerators). If a minimum regeneration solution is preferred, then the transparent option becomes the best one, with a spectrum requirement of 32 slots. Instead, if a minimum spectrum solution is desired, the fifth option is the best one, because it uses only 18 frequency slots, requiring only 2 regenerators. The opaque option (8) also uses 18 frequency slots, but requires 3 regenerators; thus, it is not Pareto optimal. We can observe that option (3) also belongs to the Pareto frontier. Indeed, it can be considered as an intermediate option that balances frequency slots $S(r) = 20$ and number of regenerators $T(r) = 1$. Our proposed provisioning strategies always select a solution from the Pareto set, avoiding dominated solutions.

The optimal provisioning routine is illustrated in Algorithm 1. Given a new demand, a path p is computed. In our implementation we use the shortest path for simplicity. Based on this path, a set $Q(p)$ of all intermediate nodes with at least 2 unused transponders is computed. Then, a subset $Q'(p) \subset Q(p)$ is built by selecting at most ρ nodes from $Q(p)$ randomly with uniform distribution. Finally, all translucent options in $R(p)$ are computed by combining the nodes in $Q'(p)$; then, blocked solutions are discarded, resulting in subset $R'(p)$. If $R'(p)$ is not empty, then the Pareto set $P(p)$ is obtained from $R'(p)$ discarding not Pareto optimal options in terms of spectrum and regeneration use. In general, the Pareto set is composed by multiple solutions, one of which should be selected for provisioning a demand over path p.

Algorithm 1: Optimal provisioning routine

```
Given a new traffic demand from a source node to a destination node;
Given an option selection routine ($selection\_strategy$), e.g.: transparent, opaque, TA, UT;
Given a spectrum assignment routine ($spectrum\_assignment$), e.g.: First-Fit, Last-Fit;
Compute a path $p$ from source to destination;
$Q(p) \leftarrow \emptyset$;
for n in intermediate nodes of $p$ do
    if n has at least 2 unused transponders then
        append n to $Q(p)$;
$Q'(p) \leftarrow \emptyset$;
for i from 0 to $max\rho, |Q(p)|$ do
    randomly select n from $Q(p)$ with uniform distribution;
    append $n$ to $Q'(p)$;
    remove $n$ from $Q(p)$;
Compute regeneration options set $R(p)$ as the set of all possible combinations of $Q'(p)$ including the empty one;
$R'(p) \leftarrow \emptyset$ ;
for r in $R(p)$ do
    if $spectrum\_assignment(demand, p, r)$ succeeds then
        append $r$ to $R'(p)$;
if $R'(p)$ is empty then
    demand blocked;
    end;
else
    compute the Pareto set $P(p)$ from $R'(p)$;
    $r_{opt} \leftarrow selection\_strategy(P(p))$;
    demand accepted with option $r_{opt}$;
    end;
```

Different selection strategies can be adopted. We classify them in two categories. On the one hand, those strategies that always adopt the same criterion, i.e., minimize the use of spectrum resources or the number of regenerators along the path. We refer to them as *resource-agnostic* strategies, because they do not assess the actual availability of all resources along the path. On the other hand, we name *resource-aware* the strategies that consider available resources along the path to select a solution. We claim that these strategies can reduce the blocking probability, as they tend to consume network resources (spectrum and transponders) so as to avoid that one resource becomes a dominant source of blocking. In other words, it would be useless to always save spectrum resources (by promoting regeneration) if transponders become unavailable and hence the main blocking source. Similarly, it would be useless to save on regeneration cost if spectrum becomes fragmented or scarce, resulting in the main blocking cause. In this context, resource-aware provisioning strategies have a great potential to efficiently exploit network resources, reducing the blocking probability and increasing the transferred traffic. In Sec. 5 we propose two strategies that demonstrate the benefits of resource-aware strategies. These strategies are based on Algorithm 1 but differ in the option selection routine ($selection\_strategy$).

4. System Model

To determine all possible solutions in $R(p)$ for provisioning a traffic demand over a path p, we introduce the main assumptions and the necessary background. In particular, we discuss how to determine for each $r \in R(p)$ the associated cost in terms of spectrum $S(r)$ when $T(r)$ additional transponders are used for regeneration purposes. We briefly describe the general network model as well as the ROADM one, and then provide a more detailed analysis of the transponder model which has a direct impact on the spectrum cost $S(r)$.

4.1. Network Model

The network is modelled as a graph $G = (V, E)$, where V represents the set of ROADM nodes and E the set of optical links. We assume that each link $e \in E$ has 320 frequency slots of 12.5 *GHz* which occupy the full C-band (4 *THz*), following the flex-grid definition [1]. We denote by K the set of traffic demands, each demand $k \in K$ is represented as a triplet $k = (s, d, c)$ where $s, d \in V$ are the source and destination nodes respectively, and c is the demand rate in *Gb/s*. Traffic demands arrive one at a time and are provisioned if enough network resources are available; otherwise, the demand is blocked.

To provision a traffic demand over a path $p \subset E$, a feasible solution r must be selected from $R(p)$. In general, a solution $r \in R(p)$ is made up of $|r|$ subpaths of p, $r = \{q_1, q_2, ...q_{|r|}\}, \quad q_i \subset p$, where each subpath can have its own modulation format, number of carriers and frequency slots. The amount of solutions depends on the hop count (i.e., fiber links) of path p as $|R(p)| = 2^{|p|-2}$, which includes not only all translucent paths but also the transparent case ($|r| = 1$) and the opaque one ($|r| = |p|$).

If the demand is provisioned over a transparent option r over path p, r is a set of only one transparent subpath $r = p$. The amount of frequency slots needed for the transparent transmission over path p is given by $F(p)$. Since $|p|$ links are traversed over path p and each link requires the same amount of $F(p)$ slots, the resulting cost in terms of spectrum is given by $S(p) = |p| \quad F(p)$ frequency slots. No additional transponders are required at intermediate nodes. Instead, if the demand is provisioned by means of a non-transparent path r, additional transponders are required at $|r| - 1$ intermediate nodes for regeneration purposes. This introduces a regeneration cost given by $T(r) = |r| - 1$. The translucent path r also requires $S(q_i)$ frequency slots on each sub-path q_i, resulting in a total spectrum cost given by $S(r) = \sum_{i=1}^{|r|} S(q_i)$ frequency slots. As the modulation format and the number of carriers may change from one subpath to other, the number of slots per link $F(q_i)$ may be different for each q_i. In general, since higher modulations formats can be used in translucent paths with respect to transparent ones, the spectrum cost can decrease as additional transponders are used for regeneration.

4.2. ROADM Model

We consider a colorless-directionless-contentionless (CDC) ROADM architecture with support of flex-grid spectrum, as described in [4]. This architecture supports optical pass-through with any amount of 12.5*GHz* frequency slots as well as super-channel add-drop capabilities. Each ROADM node is equipped with a set of transponders which can be used for provisioning demands either if the demand has as a source or destination the node itself or if regeneration is used at the node. Indeed, regeneration is supported at the node by a pair of transponders in a back-to-back configuration.

As discussed earlier in this paper, we can analyze two scenarios that account for how regeneration is considered. One scenario considers full regeneration capacity, where nodes can are equipped with enough transponders in order to fully support all opaque transmissions. Hence, in this scenario the problem is reduced to regeneration cost minimization. Instead, the other scenario assumes bounded regeneration, where each node is equipped with a limited number of transponders. Thus, the problem becomes an assignment problem which aims at optimizing the use of these transponders to reduce the blocking probability.

4.3. Transponder Model

We consider an elastic transponder model supporting a maximum capacity C expressed in *Gbit/s*. The transponder modulates a set of M optical carriers at a baud rate of

B *GBaud*, which includes an overhead of $H\%$ for forward error correction (FEC).

The estimation of the transmission reach of optical superchannels is still an open issue which has been addressed by both simulation and analytical models. Most works related to planning strategies rely on pre-computed tables that assign a transmission reach to each modulation format but typically consider only linear effects, e.g.: [8, 16, 4, 28, 29]. Instead, we estimate the transmission reach by means of the gaussian noise (GN) model [30], which also takes into account non-linear effects, resulting in a more detailed model of the physical layer. In this model, both the ASE noise introduced by optical amplifiers and the excess noise-like disturbance generated by non-linear fiber propagation called non-linear interference (NLI) are approximated as additive gaussian noise. More details on this model can be found in [30]. To limit the complexity, we opted for considering the worst-case estimate of NLI, i.e., the one referred to the center channel in case of full spectral load. On the other hand, the NLI generation is a relatively narrow-band phenomenon, as it is roughly limited to about 1 THz of bandwidth as shown in [30]. Therefore, supposing a 4 THz bandwidth for the optical fiber, with a spectral load larger-equal 25% the considered worst-case is already reached. Besides, we use the same physical layer model for all strategies analyzed in this paper to result in a fair comparison of them.

We consider six modulation formats: BPSK, QPSK, (8, 16, 32, 64)-QAM with their spectral efficiency $\eta \in \{1, 2, 3, 4, 5, 6\}$ respectively. The bit rate for each optical carrier is given by $2B\eta/(1 + H/100)$ when considering Polarization-division multiplexing (PDM). The bit rate C of the transponder can be defined as the product of all carriers rates, as shown in Eq. 1.

$$C = \frac{M \quad 2 \quad B \quad \eta}{1 + H/100} \tag{1}$$

To transport a demand $k = (s, d, c)$ the capacity of the transponder must be at least equal to c; thus, $C \geq c$. We assume that i) the number of carriers M is large enough to always satisfy the required rate, and ii) transponders support variable baud rate, with $B \leq B_{max}$ expressed in *GBaud*, where B is configured as the minimal required baud rate to satisfy the traffic demand ($C = c$) given a modulation format and the number of required carriers. To this end, each transponder, regardless being a user transceiver or a regenerator, is configured according to the following procedure:

Given a target reach (i.e., path length) and a traffic demand with rate c

1. Find the highest modulation format (spectral efficiency η) supported by the path length using the GN model [30].
2. Determine the minimum number of carriers needed for rate c
$$M = \left\lceil \frac{c \quad (1 + H/100)}{2 \quad B_{max} \quad \eta} \right\rceil \tag{2}$$
3. Update the required baud rate $B < B_{max}$ to match rate c
$$B = \frac{c \quad (1 + H/100)}{2 \quad M \quad \eta} \tag{3}$$
4. Compute the number of required frequency slots as
$$F = M \quad \left\lceil \frac{B}{12.5\,GHz} \right\rceil \tag{4}$$

Once the best modulation format is selected for a given path, both the number of carriers M and their required baudrate B to match the traffic rate c are assigned. This aims at always selecting the most efficient modulation level and the minimum number of carriers given the maximum baudrate B_{max} that the carriers can support.

In Fig. 1, where a 400 *Gb/s* service demand needs to be provisioned ($c = 400$ *Gb/s*), we can analyze the case of the transparent option $r = \{p\}$. Based on the transmission length of 2750 *km*, using the GN model [30] we can find that the highest achievable modulation format is 8QAM, with a spectral efficiency of $\eta = 3$. Assuming a maximum baud rate of $B_{max} = 50$ *GBaud* and a FEC overhead of 25% ($H = 25$), we can determine the number of required carriers using Eq.2

$$M = \left\lceil \frac{400\,Gb/s \quad (1 + 25/100)}{2 \quad 50\,GBaud \quad 3} \right\rceil = \lceil 1.66 \rceil = 2$$

. Once the number of carriers is found, the required baud rate can be computed using Eq.3

$$B = \frac{400\,Gb/s \quad (1 + 25/100)}{2 \quad 2 \quad 3} = 41.66\,GBaud$$

. Finally, the required frequency slots can be found using Eq. 4

$$F = 2 \left\lceil \frac{41.66\,GBaud}{12.5 GHz} \right\rceil = 2 \lceil 3.33 \rceil = 8$$

. Since each carrier requires 4 frequency slots, the total amount results in 8 slots. Hence, the spectrum cost required by the transparent option over 4 fiber links is the sum of the frequency slots required at each one.

$$S(r) = \sum_{i=1}^{|r|} S(q_i) = S(p) = (4)(F) = (4)(8) = 32$$

.

5. Candidate Selection

Given an optimal set $P(p)$ of candidate solutions for provisioning a traffic demand over path p, one of them needs to be selected. As earlier discussed, we argue that this decision needs to be aware of the available resources along path p. In this section, we describe two different strategies that can be used for this purpose.

5.1. Threshold Aware

In [14], we proposed an algorithm which selects the best solution given a spectrum budget. We refer to this strategy as *threshold aware* (TA), as a threshold α_S on frequency slots is used to determine the best candidate. The value of α_S can vary for each demand or depend on its actual rate, representing a target spectrum efficiency for serving demands. The main idea behind this strategy is that regeneration is a cost that needs to be minimized, but within a spectrum budget derived from α_S. If all available solutions along the path require more spectrum resources than α_S, regeneration can be used to minimize the spectrum cost. In other words, α_S behaves as a decision value to either solve a minimization problem in terms of a regeneration cost (if there exist at least one solution that requires less spectrum than α_S) or, a spectrum cost (if such solutions do not exist). This selection process is described in detail in Algorithm 2.

Algorithm 2: Threshold aware selection strategy

```
Given the Pareto set for path $p$, $P(p)$;
Given a threshold $\alpha$;
$P'(p) \leftarrow \{r \in P(p) \mid S(r) \leq \alpha_s\}$;
if $P'(p)$ is not empty then
    find $r \mid T(r)$ is $min\{T(r) | r \in P'(p)\}$;
    option $\leftarrow r$;
else
    find $r \mid S(r)$ is $min\{S(r) | r \in P(p)\}$;
    option $\leftarrow r$;
return option;
```

In Fig.3-A we show the Pareto optimal solutions in terms of spectrum and regeneration costs. The transparent solution is on the right hand side of the chart, and it requires S_{max} frequency slots and 0 regenerators. Instead, in the left hand side, we have the opaque solution with T_{max} regenerators and S_{min} frequency slots. Preferred options are those that require less frequency slots than α_S, and are highlighted inside a green area. In Fig.3-B we show the case in which, for a given a threshold, there are several preferred options. In that case, the solution that requires less generation is selected from the preferred set (green arrow). On the other hand, in Fig.3-C we show the case in which there are no options that require less frequency slots than α_S. When this happens, the strategy selects the solution with less spectrum use (green arrow).

5.2. Utilization Aware

The TA strategy requires to assign an α_S value for each demand, which can be related to a target spectrum efficiency (ratio between rate and spectrum). Besides the need to define a proper value for α_S, this approach is suitable for scenarios with full regeneration capacity , and regeneration resources are enough to meet a target efficiency. However, if regeneration is bounded, then a different approach is needed.

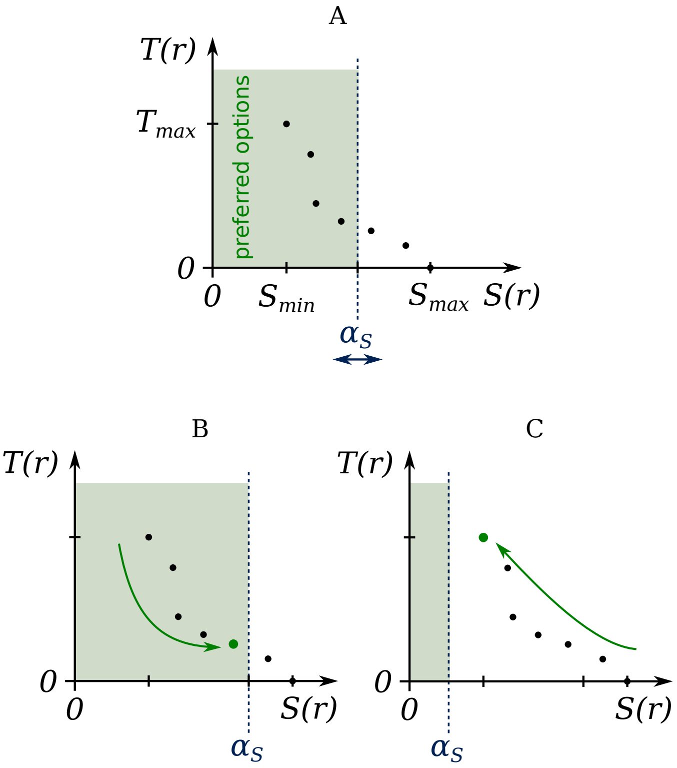

Figure 3: Threshold Aware selection strategy.

We introduce the *utilization aware* (UA) algorithm, which selects the solution that best assigns available resources to avoid the depletion of either spectrum or transponders. The UA main goal is to allocate resources to minimize the overall blocking probability and, hence, maximize the network traffic. To this end, the utilization of spectrum $U_s(p)$ and transponders $U_t(p)$ along path p is first computed to determine which resource is more critical. In both cases utilization is defined as the amount of used resources along the path with respect to existing ones. For example, $U_s(p)$ is computed as the ratio of the sum of already assigned frequency slots at each fiber link along the path over the sum of all 320 frequency slots provided by each link. While we assume that each fiber link always offers the same amount of frequency slots, the amount of available transponders at each node may vary and be proportional to the node degree (i.e., number of links the node is connected to). Hence, $U_t(p)$ is computed as the ratio of the sum of in-service transponders along the path over the sum of all available transponders at each node, which can vary for each node.

With this approach, candidate selection is always done such that provisioning has the minimum impact on the most used resource. If $U_s(p)$ is larger than $U_t(p)$, then the spectrum is considered the critical resource that needs to be saved. Hence, the candidate solution r that minimizes the spectrum cost $S(r)$ is selected. However, if $U_t(p)$ is larger than $U_s(p)$, then the candidate solution r that minimize regeneration cost $T(r)$ is chosen.

However, a straightforward implementation of this strategy may lead to unnecessary blocking conditions. Lets assume that we use an algorithm that always minimize the most used resource. When the utilization is low there is no risk in selecting near-opaque or near-transparent solutions because the available resources are abundant. Instead, when utilization becomes high in both dimensions, considering only the most used resource to select the best solution can be counterproductive, as it can result in an excessive increase on the least used resource, which can then become the most critical resource.

An alternative to select the best candidate when utilization is high, is to use a min-max approach over all candidates. This implies that, for each candidate r, the resulting utilization in terms of $U_s(r)$ and $U_t(r)$ is estimated, and its maximum value $max(U_s(r), U_t(r))$ compared with all other maximum values from candidates in the Pareto set $P(p)$. The candidate with the minimum maximum utilization is then selected. This strategy selects solutions that increase the utilization of both resources, hence, equalizing utilization. Even if this could be a wise decision when resources are scarce (i.e., highly utilized), it may not be appropriate when they are abundant (i.e., low utilization). In the latter case, it becomes more relevant to reduce the rate at which the maximum utilization increases, as demands are provisioned instead of caring about their equalization. For example, if spectrum resources are more abundant than transponder ones, then the transponder utilization rate needs to be controlled. This would favour the selection of solutions with minimum or zero (transparent) regeneration. However, as utilization increases, more balanced solutions are required to avoid the earlier depletion of one resource.

To tackle this challenge, we propose to control the candidate solution space as a function of the utilization. As one resource increase its utilization, the solution space tends to more balanced solutions in terms of spectrum and transponder cost. To this end, the UA algorithm uses two thresholds, one in terms of spectrum (α_S) and another in terms of regenerators (α_T), to define a set of preferred candidates from the Pareto set $P(p)$. If the spectrum utilization $U_s(p)$ along path p is lower than the transponder one $U_t(p)$, then the solution r inside α_S and α_T thresholds that minimizes the regeneration cost $T(r)$ is selected. Otherwise, if $U_t(p)$ is lower than $U_s(p)$, then the solution r that minimizes the spectrum cost $S(r)$ within the thresholds is chosen. If the preferred set is empty, then the selection is done over the whole Pareto set.

In Fig.4 we show the proposed policy. Each black dot represents a candidate solution. The opaque option requires T_{max} regenerators and S_{min} frequency slots while the transparent solution requires zero regenerators and S_{max} frequency slots. There are two budgets: the regenerators threshold α_T and the spectrum threshold α_S. The preferred solutions are those that are simultaneously under α_T and α_S (green area).

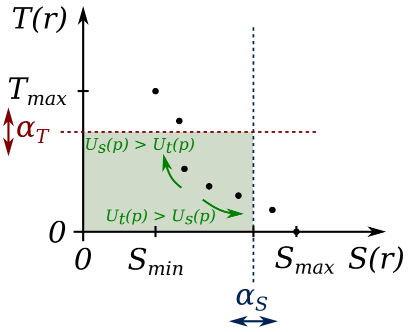

Figure 4: Two dimension bounded utilization aware option selection strategy.

The option selection is done by minimizing the most used resource over the options that are simultaneously under α_S and α_T (preferred set). If the preferred set is empty, then the optimization is done over the whole Pareto set also minimizing the most used resource.

To adapt to the changes of network state and to the demand requirements (bitrate, path length), the thresholds are dynamically computed for each new provisioning process.

A demand can be provisioned using a number of regenerators that ranges from 0 (transparent case) to T_{max}, that is the number of intermediate nodes in the selected path. For example, if a 6 hops path is selected to provision a demand, then T_{max} is going to be 4 for this demand. We propose to set a transponder threshold (α_T) that is computed for each demand from this range using the path utilization. When the path transponder utilization is low, we set the threshold to T_{max}, meaning that we prefer any solution regardless the required number of regenerators. As the utilization starts to increase, α_T decreases proportionally. So, the regenerators threshold can be computed as $\alpha_T = T_{max}(1 - U_t(p))$, where $U_t(p)$ is the transponder utilization of the selected path. In Eqn.5 we show the complete formula for α_T. We added another factor in the product $(1 - \beta)$ to ensure that the opaque solution is included when the utilization is near zero (β is a small value determined by tuning).

In a similar way, the spectrum threshold depends on the spectrum path utilization ($U_s(p)$) and the spectrum requirements of the opaque and transparent solutions (S_{min} and S_{max}) following the expression shown in Eq.6. In this case, as the spectrum requirements of a demand ranges from S_{min} (the spectrum required by the opaque solution) and S_{max} (the spectrum required by the transparent solution), the utilization is multiplied by $S_{max} - S_{min}$ and then is added to S_{min}. Again, we introduced a $(1 - \beta)$ factor multiplying the spectrum requirements to ensure that the transparent solution is included in the preferred options set when the spectrum utilization is near zero.

$$\alpha_T = T_{max}(1 - \beta)(1 - U_t(p)) \tag{5}$$

$$\alpha_s = (S_{max} - S_{min})(1 - \beta)(1 - U_s(p)) + S_{min} \tag{6}$$

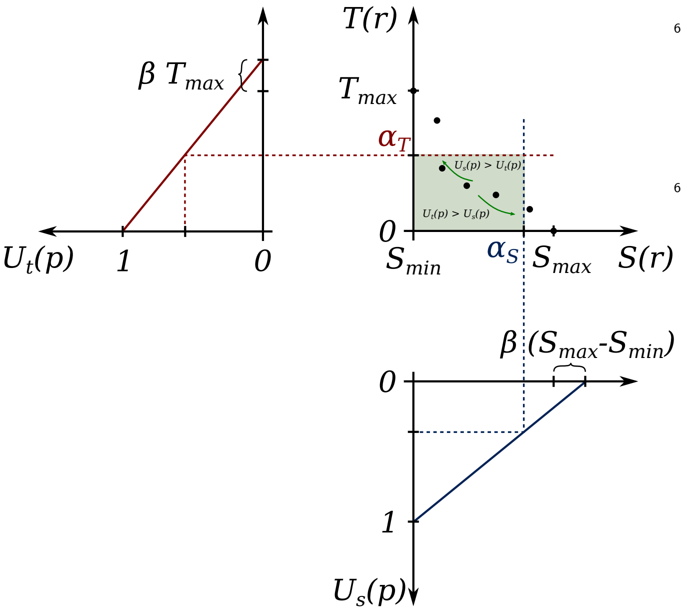

Figure 5: Utilization Aware selection strategy.

In Fig. 5 we show the proposed method for the candidate selection. Given a new demand that needs to be provisioned over a path, the Pareto set is determined. Then, both thresholds (α_S, α_T) are computed using Eq. 5 and 6. These thresholds determine the preferred set of options (green area), that is composed by all the regeneration options that require less frequency slots than α_S and less regenerators than α_T. Finally, the option that minimize the most used resource is selected from the preferred set. If the preferred set is empty, then the option is selected from the Pareto set, minimizing the most used resource.

The spectrum threshold is computed as a linear function of the spectrum utilization parametrized by the opaque and transparent solutions spectrum requirements, S_{min} and S_{max}.

The threshold ranges between S_{min} and $S_{max}+\beta(S_{max}-S_{min})$. When the utilization is small, the threshold gets a value slightly greater than S_{max} (β is a small value). Then, no bound in terms of spectrum is imposed but, as the spectrum path utilization ($U_s(p)$) increases the threshold decreases. When the utilization tends to 1, the threshold tends to S_{min}, meaning that more regenerated options are preferred. In a similar way, the regeneration threshold is computed as a linear function of the transponder path utilization parameterized by the opaque and transparent solutions regeneration requirements (T_{max}, 0). When the transponder path utilization is near zero, the threshold is over T_{max} and all solutions are included as preferred. As the utilization increase the threshold decreases, reaching zero when the utilization is one. Overall, when the network is not loaded and the utilization is small, all the possible solutions are included. As the utilization increases the preferred set becomes more and more bounded and more balanced solutions are more likely selected. As the most used resource is minimized, the spectrum and transponder utilization are better balanced for successive demands. The β parameter is a small value close to zero that ensures that the opaque and transparent solutions are included when the utilization is near zero.

The details of the option selection routine can be found in Algorithm 3. This strategy adjusts the thresholds (α_T and α_S) to the current state of the network using the spectrum and transponder path utilization as a reference.

Algorithm 3: utilization based option selection

```
Given the Pareto set for path $p$, $P(p)$;
Compute $U_t(p)$ (transponder utilization of $p$);
Compute $U_s(p)$ (spectrum utilization of $p$);
Compute $S_{max}(p)$ (frequency slots, transparent option);
Compute $S_{min}(p)$ (frequency slots, opaque option);
Compute $T_{max}(p)$ (number of regenerators, opaque option);
$\alpha_T \leftarrow (1-\beta) \times (1-U_t(p)) \times T_{max}(p)$;
$\alpha_S \leftarrow (1-\beta) \times (1-U_s(p)) \times (S_{max}(p)-S_{min}(p)) + S_{min}(p)$;
$P'(p) \leftarrow \{r \in P(p) \mid S(r) \leq \alpha_S \ and \ T(r) \leq \alpha_T\}$;
if $P'(p)$ is empty then
    $P'(p) \leftarrow P(p)$;
if $U_t(p) > U_s(p)$ then
    find $r \mid T(r)$ is $min\{T(r) | r \in P'(p)\}$;
    option $\leftarrow r$;
else
    find $r \mid S(r)$ is $min\{S(r) | r \in P'(p)\}$;
    option $\leftarrow r$;
return option;
```

If the initial solution set $R(p)$ set includes all the feasible solutions, then the complexity of our proposed algorithms becomes combinatorial. Hence, we introduce a bound on the number of intermediate nodes that are included in the base of the combinations. We call this bound ρ. Once the path is set, ρ intermediate nodes are randomly selected to form the base discarding those that do not have available transponders. Therefore, the base will have at most ρ intermediate nodes and the maximum number of regeneration options is bounded to 2^{ρ}, keeping the complexity manageable.

5.3. Resource-Agnostic Strategies

Since we have described our proposed strategies as resource-aware, we briefly discuss existing strategies that can be considered resource-agnostic. The most straightforward ones are the transparent and opaque strategy. These strategies represent trivial and opposite strategies, which try to minimize only one of the costs. On the one hand, in the *transparent* strategy no regeneration is ever used, which can result in an intensive use of spectrum resources. On the contrary, the *opaque* strategy, which uses regeneration at each hop, can save spectrum but at the expenses of

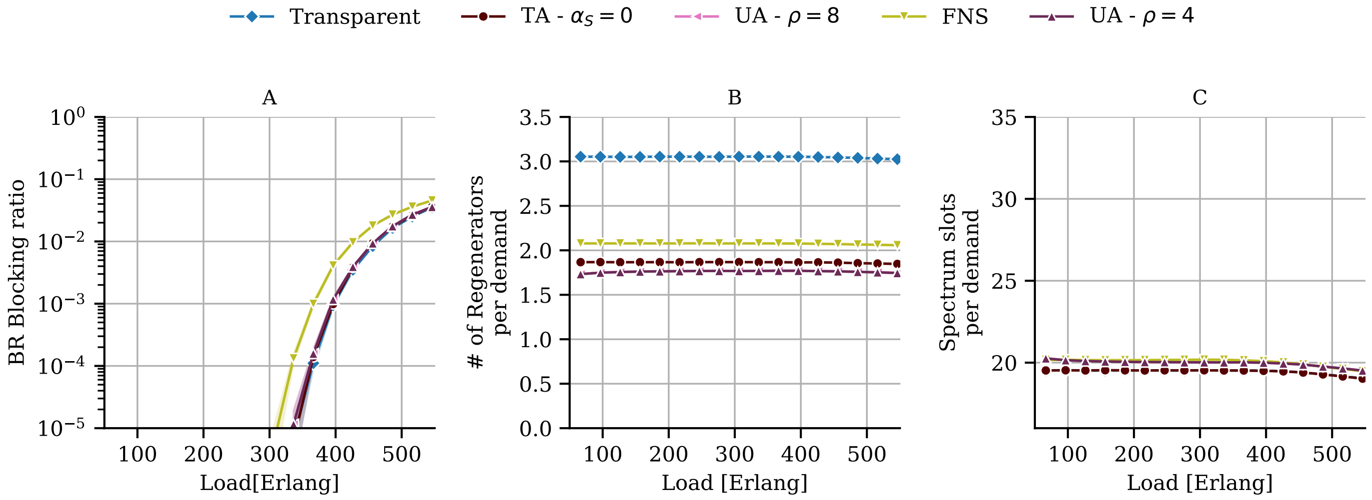

Figure 6: Simulation results for the COST266 European Topology under dynamic mixed rate traffic and 80 transponders per link per node. A) Bitrate blocking ratio vs Load, B) Number of regenerators per demand vs Load, C) frequency slots per demand vs Load.

increasing the amount of transponders required for regeneration. Both strategies can be considered as upper and lower bounds for the analysis of other strategies.

The *first longest reach* (FLR) strategy is a natural improvement of the *transparent* one, and it is proposed in [24]. Regeneration is only used if a transparent path cannot be setup. The strategy always tries to establish the longest transparent path that is possible. Given a path, the algorithm increases the transparent transmission length by one hop until it becomes blocked. If blocking occurs, a regenerator is placed as far as possible from the source. Ultimately, the behaviour is near to the transparent case but with the extra capability of overcoming some blocking situations using regeneration

On the other hand, the *first narrowest spectrum* (FNS), also proposed in [24], aims at improving the *opaque* strategy by using the narrowest spectrum as far as possible. Given a path, the algorithm increases the transparent transmission length by one hop at a time, until the amount of frequency slots per link increases. At this point, a regenerator is placed in the previous node. The algorithm tries to minimize the spectrum use in the first place and, as a second step, to minimize the number of regenerators. This algorithm finds solutions near to the opaque case, but potentially using less transponders.

6. Results

In this section we analyse simulation results to evaluate the proposed strategies and compare them with existing ones. Our main goal is to demonstrate the value of strategies that are aware of available resources. First, we consider those strategies that assume full regeneration capacity, where nodes are equipped with a large amount of transponders. Hence, these strategies aim at minimizing the required spectrum with the minimum regeneration cost. Next, we consider those strategies that are well suited for scenarios where the number of transponders per node is a limiting resource. These strategies aim at optimally assigning available resources. The bound on the amount of available transponders per node is also analyzed to better understand how this impacts the blocking probability. Finally, we present some complexity considerations of the proposed strategies and compare their execution time with respect to existing ones.

We ran simulations for three optical network topologies: the European COST266 (37 Nodes and 57 Links), the United States US24 (24 Nodes and 43 Links) and the Argentinean ARG-CORE (29 Nodes and 55 Links). These topologies can be found in net2plan format in `https://github.com/nehuengonzalez/eon_topologies`. We only report results for the European COST266 Network topology, that has 37 Nodes and 57 Links, as result trends are equivalent to those obtained for other topologies that were simulated. We assumed SMF fiber, $\alpha = 0.2$ *dB/km*, $D = 17$ *ps/(nm km)*, $\gamma = 1.2$ *1/(W km)* and EDFA optical amplification every 80 *km* with 5 *dB* of noise figure. We consider the whole C-band (4 THz) divided in 320 frequency slots of 12.5 GHz to represent the flexgrid [1]. For simplicity, we consider one empty slot as guard-band between adjacent super-channels. We consider that the amount of transponders available at each node is proportional to the node degree, for example, if we consider that each node will have 10 transponders per link a 3 degree node will have 30 transponders. An elastic transponder supporting BPSK, QPSK, (8, 16, 32, 64)-QAM ($\eta \in \{1, 2, 3, 4, 5, 6\}$), 25% FEC overhead targeting BER=1e-2 ($H = 25$), and a maximum baud rate of 50 GBaud ($B = 50$) is assumed.

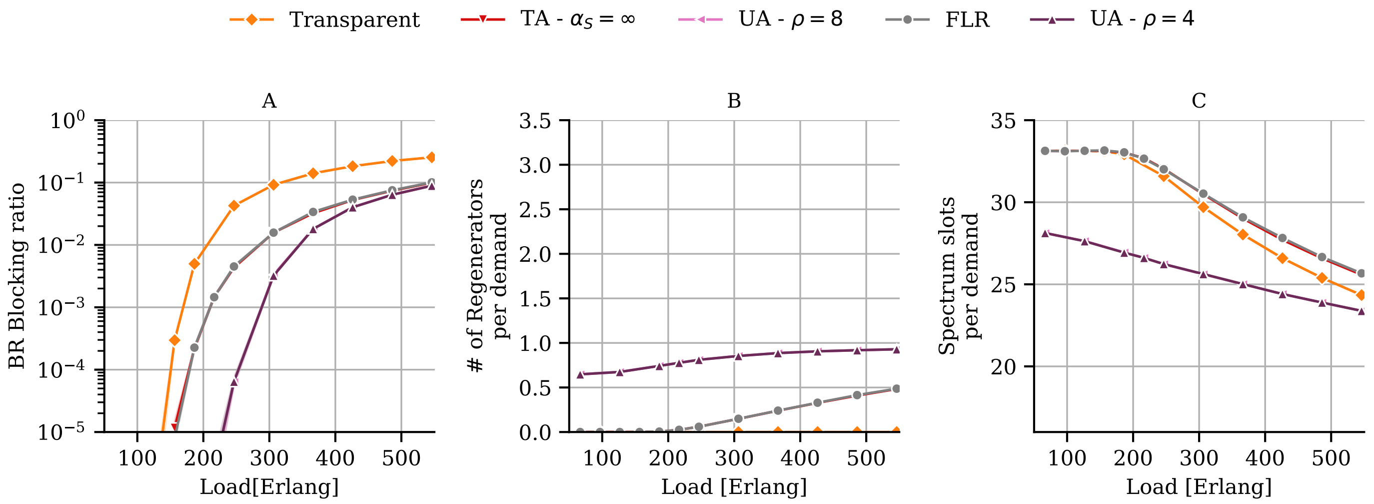

Figure 7: Simulation results for the COST266 European Topology under dynamic mixed rate traffic and 20 transponders per node per link. A) Bitrate Blocking ratio vs Load, B) Number of regenerators per demand vs Load, C) frequency slots per demand vs Load.

We consider mixed rate dynamic traffic demands of 100 *Gb/s*, 200 *Gb/s* and 400 *Gb/s*. Demand arrival and holding times follow independent Poisson processes. The source and destination node, and the demand bitrate are randomly selected using uniform distributions. All the results shown were obtained running 30 independent simulations with 30 different random seeds with 500000 demand arrivals each. We also report the confidence interval of the values as a shadow around the dots: In most of the cases the confidence interval is so close to the mean value that can not be appreciated.

6.1. Full Regeneration Capacity

Strategies that make an intensive use of transponders are well suited for scenarios where the limiting resource is the spectrum. We analyze a scenario with full regeneration capacity where each node has enough transponders to admit fully opaque transmission in all the demands, getting almost no transponder blocking. **In Fig. 6-A we show the blocking ratio versus the network load for this scenario. We consider 80 transponders per link per node to match the number of channels of 50GHz in the band C. The 50GHz bandwidth is selected to match the maximum baudrate of the transponders (50 *GBaud*).** In this scenario the limiting resource is the spectrum and the *opaque* strategy gets the best performance due to the spectrum compression obtained thanks to the regeneration. A similar performance is obtained by the *TA* strategy when $\alpha = 0$ and by the *UA*. For *TA*, such a low threshold value tends to preferentially select a solution that has the minimum spectrum cost but with the minimum number of regenerators. For *UA*, the assumption of large amount of transponders at each node always results in low transponder cost $T(p)$ for any path p. Hence, the spectrum is always assumed to be the critical resource and solutions that have minimum impact on it are selected. Note also that we considered two values for ρ, which set a limit on the number of nodes where regeneration is analyzed. The larger its value, the more candidate solutions are considered. Since there is no significant difference on the performance for both values, we can argue that the *UA* complexity can be limited with no actual penalty. Later we will evaluate the complexity of this strategy, but we can anticipate that the best performance can be achieved with low complexity. Finally, we observe that the *FNS* strategy has the worst performance. For low loads, this strategy gets blocking ratios even ten times worse than the others, and for higher loads that difference reduces to two times.

In Fig. 6-B we show the number of regenerators per demand versus the load. The *opaque* strategy requires a very high number of regenerators per demand, slightly more than three on average. independently of the load. In this scenario, with a large amount of transponders per node, the number of regenerators can be thought as a relative cost. For example, three regenerators per demand represent a three times increment in the transponder cost compared to the transparent transmission. All other strategies can reduce this cost, even for similar blocking performance. The *opaque* strategy uses regeneration at every intermediate node, regardless of the impact of that regeneration on the spectrum use. On the other hand, the *TA*-$\alpha_S = 0$ and the *UA* strategies select the best option from the Pareto set, then a candidate with less regeneration and minimal impact on the spectrum is always favoured. This means that using the *resource aware* strategies the best performance achievable can be met even if we reduce the amount of transponders per link per node in a 30% compared with the requirements of the opaque strategy.

Both *TA*-$\alpha_S = 0$ and *UA* strategies not only perform

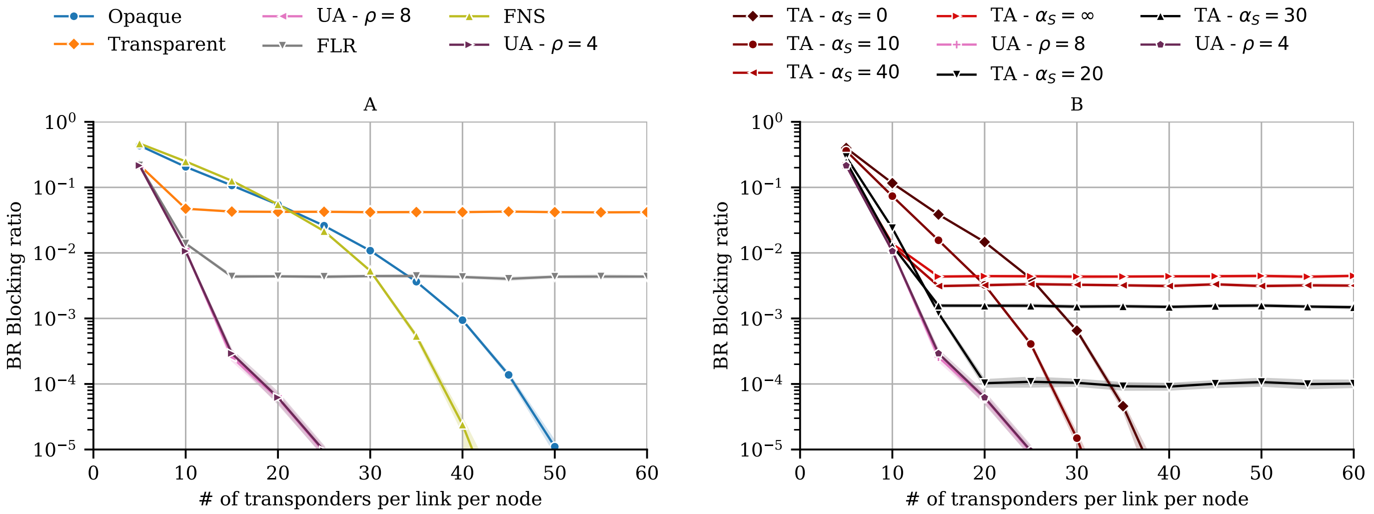

Figure 8: Simulation results for the COST266 European Topology under dynamic mixed rate at 250 Erlang. Bitrate Blocking ratio vs transponders per link per node

better in terms of blocking than the *FNS* one, but use less regeneration. Furthermore, these two strategies use less spectrum per demand, as shown in Fig. 6-C. Overall, when the limiting resource is spectrum, the *TA*-$\alpha_S = 0$ and *UA* strategies have the best performance in all three dimensions: blocking ratio, number of regenerators and frequency slots per demand.

6.2. Partial Regeneration Capacity

A more realistic examined scenario is when both resources are scarce. In such scenario, strategies that make an intensive use of spectrum resources may be more suitable. In contrast to Sec. 6.1, nodes are equipped with much less transponders. In Fig. 6 we show results for the same topology but with 20 transponders per link per node. We consider the *transparent*, *TA*-$\alpha_S = \infty$, *UA* and the *FLR* strategies.

In Fig. 7-A we show the blocking ratio versus network load. As it can be seen, in this scenario the blocking probabilities are much higher due to the lack of transponders. These strategies use a small amount of regenerators per demand (Fig. 7-B) compared to the previous strategies (Fig. 6-B). This behaviour has a direct impact on the number of slots per demand, as shown in Fig. 7-C. These strategies require more than 23 frequency slots per demand at any load, while the previous strategies require less than 20. This increment on the spectrum requirements is reflected on the increment of the blocking ratios, since the spectrum resources become scarce at lower network loads.

The *transparent* strategy does not allow any regeneration at all, and as a consequence it has the worse blocking ratio performance. The *FLR* strategy uses regeneration whenever the transparent solution is blocked, enabling a blocking probability gain of hundred of times at low load conditions and of two times at very high load conditions. The *TA* algorithm with an α_S set to infinity always tries to minimize the use of transponders, then it has a very similar behaviour to *FLR*. In both cases, the number of transponders grows with the increment of the load. As the load increases, the probability of getting blocked by spectrum continuity and contiguity problems also increases. Thus, both strategies use regeneration to overcome this blocking situations. Finally, the *UA* try to balance the spectrum use and transponders use based on the basis of spectrum and transponders path utilization. This balanced solution enables a further reduction on the blocking probabilities, with gains that are over thousand of times at low load conditions and up to two times in very high load conditions.

6.3. Regeneration Capacity Sensitivity

In Fig. 8 we show simulation results for the same topology at a single load of 250 Erlang for multiple regeneration capacities in terms of transponder bounds. We select this load because it gives a wide range of observable blocking probabilities, but similar behaviours are obtained for other load conditions. We divide the results in two sets to provide clarity. In Fig. 8-A we compare the *UA* strategy with *transparent*, *opaque*, *FLR* and *FNS* ones. In Fig. 8-B we compare the *UA* strategy with the *TA* one.

In Fig. 8-A the *UA* always perform better than the other strategies. When the transponder bound is small (below 10 transponders per link per node), the results obtained with the *UA* algorithm overlaps with those obtained with the *FLR*. Indeed, in such scenario transponders are always more scarce than frequency slots as a consequence, the best solution is always the one that minimizes the use of regeneration. This behaviour is also reflected by the fact that, when the transponder bound is 5 transponders per link per node, no strategy performs better than the *transparent*. There is a point from which the *transparent*

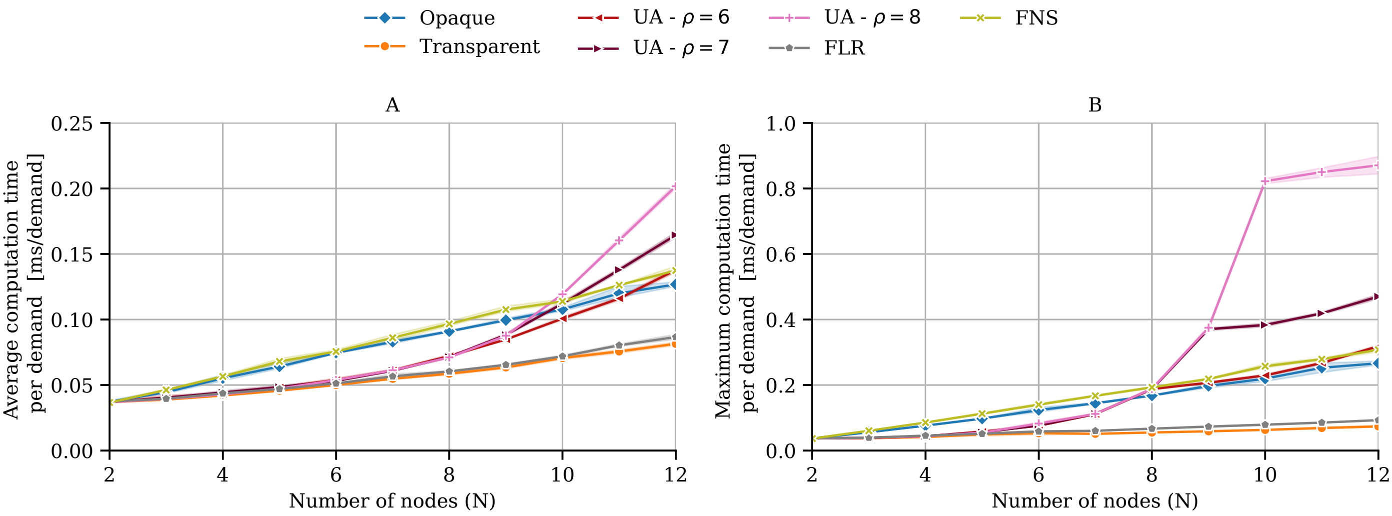

Figure 9: Computation time per demand for a linear topology with N nodes. A) Average computation time per demand vs. Number of nodes, B) Maximum computation time per demand vs. Number of nodes.

and *FLR* algorithms can no longer obtain better results, since they will not use extra regeneration even if spare transponders are available, As the transponder bound increases, those strategies that make a more intensive use of transponder start improving their performance. For these transponder bounds, the *UA* algorithm always shows the best performance, because it always tries to use the available resources in both dimensions.

In Fig. 8-B we compare the *UA* algorithm with the *TA* algorithm with a range of thresholds $\alpha_S = 0, 10, 20, 30, 40, \infty$. Observe that, for every transponder bound, there exist an α_S such that the *TA* algorithm perform almost as well as the *UA*. If we had an oracle to determine the best α_S for each scenario we could always obtain good results. However, being the definition of an optimal α infeasible, this strategy is not practical. The *UA* strategy aims to use the best resource combination at every step, taking into account the network state and the required resources. This is why it always permits to obtain good results regardless the transponder bound.

6.4. Complexity Analysis

The most computationally intensive part of all these algorithms is the spectrum allocation step. All the algorithms are using the same spectrum allocation subroutine (i.e., First-Fit scheme). Then the complexity of the different algorithms depends on the number of times the spectrum allocation subroutine is called. This subroutine evaluates both spectrum and transponder availability along a path or subpath. If there is no spectrum nor transponder blocking, it determines the best spectrum allocation.

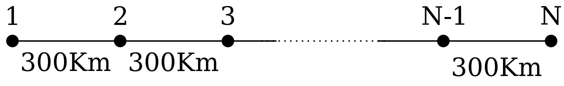

Figure 10: N nodes linear topology.

The worst case complexity of the *UA* algorithm is dominated by the number of combinatorial options needed to be evaluated. Let assume that the complexity of the spectrum allocation is constant. Let H be the number of hops in the path, then the *UA* algorithm complexity is $O(2^{H-1}H)$, where 2^{H-1} is the number of regeneration options and H the number of calls to the spectrum allocation procedure per evaluated option. As the number of nodes included in the base for the combination is bounded by ρ, then the complexity is also bounded by $O(2^{\rho}H)$.

Nevertheless, the worst case complexity of the *UA* strategy is very pessimistic. In practice, the actual number of candidate solutions that are evaluated might be much lower, because some solutions may be blocked by lack of spectrum or transponder resources. The requirements of each option can be precomputed and those options that do not have enough resources are discarded, with almost no computational cost. Besides, the number of calls to the spectrum allocation is commonly much lower than the worst case H. There is only one option with $H-1$ regenerators. The average number of regenerators in a random selection is a number between 0 and $H-1$.

In contrast, the *FLR* and the *FNS* worst case complexity is lower, being equal to $O(2H)$. In the case of *FLR* and *FNS*, the best case computation time follows $O(H)$. Therefore, the average computation time will be between $O(H)$ and $O(2H)$, With the *FLR* near to $O(H)$ and *FNS* near to $O(2H)$.

In all the previous simulations, the computation time per demand in $ms/demand$ is comparable for all strategies, including *UA*. Since the computation time mainly depends on the number of hops of the paths, we analyze its actual performance on a linear topology, as shown in Fig. 10 with N nodes and a constant hop length of 300 km. We run simulations for different number of nodes N with mixed rate traffic (100G, 200G and 400G uniformly distributed) for an homogeneous traffic matrix. We computed the average computation time per demand in *ms/demand* and the maximum computation time per demand in *ms/demand* for each N, at a load that gives a blocking ratio between 0.01 and 0.05. Each dot is the result of averaging over 200 simulations with 50000 demand arrivals each. As it can be seen in Fig. 9-A, the average normalized computation times for all the strategies are in the order of magnitude regardless of N. When the number of hops of the topology is over ten, the UA starts to require higher computation time than the other strategies.

The effect of ρ can be observed in Fig. 9-B, where we show the maximum computation time per demand: For $\rho = 8$, the maximum computation time grows exponentially until $N = \rho + 2 = 10$, then the computation time growth becomes linear. Something similar happens with $\rho = 7$: The computation time starts to behave linearly after $N = 9$ and after $N = 8$ for $\rho = 6$. Including this limit, we allow the algorithm to get the global optimum when the size of the solution field is manageable and to get a suboptimum when the size is not manageable. As earlier discussed, we included a parameter ρ in the UA strategy that represents the maximum number of intermediate nodes to be included in the computation routine. In other words, if a path has more than $\rho + 1$ hops, only ρ intermediate nodes will be selected as candidates for regeneration. Thus, the number of evaluated solutions is bounded by 2^{ρ}. If, for example, we set $\rho = 10$, all the demands with paths with at most 11 hops will be provisioned using the full set of regeneration options, whereas demands with paths with more than 11 hops will be provisioned using only 1024 options. This bound is used in the third step of the flow shown in Algorithm 1, the regeneration options computation step. When the number of intermediate nodes of the path is greater than ρ, the maximum number of regeneration nodes are randomly selected according to a uniform distribution.

7. Conclusions

In this work we analyzed the trade-off between regeneration and spectrum costs when provisioning lightpaths in translucent EONs. Existing strategies aimed at minimizing either one of these costs, instead we focused on a joint optimization.To this end, we modeled the problem as a two-dimensional resource assignment problem and proposed two different provisioning strategies. The first one, named *threshold-aware* (*TA*), is well suited for scenarios with full regeneration capacity. The second one, named *utilization-aware* (*UA*), considers scenarios where regeneration capabilities at nodes are bounded.

We reported simulation results that showed that proposed *resource-aware* strategies can enhance the performance in both scenarios in terms of blocking ratio and cost. In the scenario with full regeneration capacity, where the number of regenerators per demand can be considered as a cost, both the *TA* and *UA* strategies outperform the best known algorithms achieving the same performance of the opaque provisioning strategy in terms of blocking ratio while reducing the regeneration cost by more than 30%. In the general scenario with scarce spectrum and transponders, where the number of regenerators per demand cannot be considered as a cost, the *UA* algorithm enables a more efficient use of spectrum and transponder resources, reducing the blocking ratio by thousand of times with respect to the transparent provisioning strategy and by hundred of times if compared to the best known algorithms and the *TA* algorithm.

The ultimate blocking gain depends on the overall number of available transponders but the *UA* algorithm always ensures the lowest blocking ratio compared to the rest of the alternatives. When the number of available transponders is very low, no strategy can perform better than the transparent case; similarly, when the number of available transponders tends to infinity no strategy can perform better than the opaque one.

Finally, we showed that the computation time per demand is comparable among all strategies when the path lengths are below 12 hops. We also showed that when including a limit to the number of evaluated combinations of regeneration points, the *TA* and the *UA* complexity can be bounded with almost no practical impact on performance.

In summary, these results show that *resource-aware* strategies can better exploit the flexibility of EONs and improve performance by reducing the blocking probability. Even if our work only considers spectrum and transponders as the available resources to be optimized, other resources could be considered in future studies. In particular, the use of sliceable bandwidth variable Transponders (SBVTs) introduces more flexibility in the resource assignment problem as a transponder could be shared among different lightpaths. Besides, the spectrum cost could be further extended when considering space division multiplexing (SDM) EONs, so as they add the fiber cores or propagation modes in the decision process. In this context, we foresee that provisioning strategies need to consider the trade-off among all available resources and we believe that this work contributes to this vision and provides a valid methodology to tackle these challenges.

References

[1] Itu-t recommendation “g.694.1 : Spectral grids for wdm applications: Dwdm frequency grid”, https://www.itu.int/rec/T-REC-G.694.1/en.

[2] M. Jinno, H. Takara, B. Kozicki, Y. Tsukishima, Y. Sone, S. Matsuoka, Spectrum-efficient and scalable elastic optical path network: architecture, benefits, and enabling technologies, IEEE Communications Magazine 47 (11).

[3] O. Gerstel, M. Jinno, A. Lord, S. B. Yoo, Elastic optical networking: A new dawn for the optical layer?, IEEE Communications Magazine 50 (2).

[4] B. C. Chatterjee, N. Sarma, E. Oki, Routing and spectrum

allocation in elastic optical networks: A tutorial, IEEE Communications Surveys & Tutorials 17 (3) (2015) 1776–1800.

[5] I. Tomkos, S. Azodolmolky, J. Sole-Pareta, D. Careglio, E. Palkopoulou, A tutorial on the flexible optical networking paradigm: State of the art, trends, and research challenges, Proceedings of the IEEE 102 (9) (2014) 1317–1337.

[6] J. M. Simmons, Optical network design and planning, Springer, 2014.

[7] M. Klinkowski, K. Walkowiak, On performance gains of flexible regeneration and modulation conversion in translucent elastic optical networks with superchannel transmission, Journal of lightwave technology 34 (23) (2016) 5485–5495.

[8] Klinkowski, Miroslaw and Walkowiak, Krzysztof, Performance analysis of flexible regeneration and modulation conversion in elastic optical networks, in: Optical Fiber Communication Conference, Optical Society of America, 2017, pp. Th2A–13.

[9] X. Wang, M. Brandt-Pearce, S. Subramaniam, Impact of wavelength and modulation conversion on translucent elastic optical networks using milp, Journal of Optical Communications and Networking 7 (7) (2015) 644–655.

[10] M. Klinkowski, On the effect of regenerator placement on spectrum usage in translucent elastic optical networks, in: Transparent Optical Networks (ICTON), 2012 14th International Conference on, IEEE, 2012, pp. 1–6.

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

[12] D. A. Chaves, M. A. Cavalcante, H. A. Pereira, R. C. Almeida, A case study of regenerator placement and regenerator assignment in dynamic translucent elastic optical networks, in: Transparent Optical Networks (ICTON), 2016 18th International Conference on, IEEE, 2016, pp. 1–4.

[13] J. M. Finochietto, M. Garrich, A. Bianco, On provisioning strategies in translucent elastic optical networks with flexible regeneration and superchannel transmission, in: High Performance Switching and Routing (HPSR), 2017 IEEE 18th International Conference on, IEEE, 2017, pp. 1–6.

[14] N. Gonzalez-Montoro, J. M. Finochietto, A. Bianco, Optimal provisioning strategies for translucent elastic optical networks, in: 2018 IEEE Global Communications Conference (GLOBECOM), IEEE, 2018, pp. 1–7.

[15] E. F. da Silva, R. C. Almeida, H. A. Pereira, D. A. Chaves, Assessment of novel regenerator assignment strategies in dynamic translucent elastic optical networks, Photonic Network Communications 39 (1) (2020) 54–69.

[16] K. Walkowiak, M. Klinkowski, A. Włodarczyk, A. Kasprzak, Predeployment of transponders for dynamic lightpath provisioning in translucent spectrally–spatially flexible optical networks, Applied Sciences 10 (8) (2020) 2802.

[17] S. Azodolmolky, M. Klinkowski, E. Marin, D. Careglio, J. S. Pareta, I. Tomkos, A survey on physical layer impairments aware routing and wavelength assignment algorithms in optical networks, Computer Networks 53 (7) (2009) 926–944.

[18] S. Pachnicke, T. Paschenda, P. M. Krummrich, Physical impairment based regenerator placement and routing in translucent optical networks, in: Optical Fiber Communication Conference, Optical Society of America, 2008, p. OWA2.

[19] W. Xie, J. P. Jue, X. Wang, Q. Zhang, Q. She, P. Palacharla, M. Sekiya, Regenerator site selection for mixed line rate optical networks with flexible routing, in: Optical Network Design and Modeling (ONDM), 2012 16th International Conference on, IEEE, 2012, pp. 1–6.

[20] W. Xie, J. P. Jue, X. Wang, Q. Zhang, Q. She, P. Palacharla, M. Sekiya, Cost-optimized design of flexible-grid optical networks considering regenerator site selection, in: Global Communications Conference (GLOBECOM), 2013 IEEE, IEEE, 2013, pp. 2358–2363.

[21] I. Cerutti, F. Martinelli, N. Sambo, F. Cugini, P. Castoldi, Trading regeneration and spectrum utilization in code-rate adaptive flexi-grid networks, Journal of Lightwave Technology 32 (23) (2014) 3894–3901.

[22] M. Aibin, K. Walkowiak, Adaptive modulation and regenerator-aware dynamic routing algorithm in elastic optical networks, in: Communications (ICC), 2015 IEEE International Conference on, IEEE, 2015, pp. 5138–5143.

[23] A. Fallahpour, H. Beyranvand, S. A. Nezamalhosseini, J. A. Salehi, Energy efficient routing and spectrum assignment with regenerator placement in elastic optical networks, Journal of Lightwave Technology 32 (10) (2014) 2019–2027.

[24] D. A. Chaves, E. F. da Silva, C. J. Bastos-Filho, H. A. Pereira, R. C. Almeida, Heuristic algorithms for regenerator assignment in dynamic translucent elastic optical networks, in: Transparent Optical Networks (ICTON), 2015 17th International Conference on, IEEE, 2015, pp. 1–4.

[25] N. Tokas, A. Kumar, K. Kaur, Optimum regenerator placement in wdm optical network, International Journal of Engineering Science 10839.

[26] H. Guo, Y. Li, L. Li, G. Shen, Adaptive modulation and regeneration-aware routing and spectrum assignment in sbpp-based elastic optical networks, IEEE Photonics Journal 9 (2) (2017) 1–15.

[27] N. Gonzalez-Montoro, J. M. Finochietto, A. Bianco, Translucent provisioning in elastic optical networks with sliceable bandwidth variable transponders, in: 2018 IEEE Global Communications Conference (GLOBECOM), IEEE, 2018, pp. 1–6.

[28] J. Zhao, L. Yan, H. Wymeersch, E. Agrell, Code rate optimization in elastic optical networks, in: 2015 European Conference on Optical Communication (ECOC), IEEE, 2015, pp. 1–3.

[29] X. Luo, Y. Zhao, X. Chen, L. Wang, M. Zhang, J. Zhang, Y. Ji, H. Wang, T. Wang, Manycast routing, modulation level and spectrum assignment over elastic optical networks, Optical Fiber Technology 36 (2017) 317–326.

[30] P. Poggiolini, The gn model of non-linear propagation in uncompensated coherent optical systems, Journal of Lightwave Technology 30 (24) (2012) 3857–3879.