A survey on physical layer impairments aware routing and wavelength assignment algorithms in optical networks

Siamak Azodolmolkya,b,*, Miroslaw Klinkowskia,c, Eva Marina, Davide Careglio a, Josep Solé Pareta a, Ioannis Tomkos b

a Technical University of Catalonia (UPC), C/Jordi Girona, 1-3, 08034 Barcelona, Catalonia, Spain
b Athens Information Technology (AIT), Markopoulo Ave., 19002 Peania, Athens, Greece
c National Institute of Telecommunications (NIT), 1 Szachowa Street, 04-894 Warsaw, Poland

1. Introduction

During the past couple of years, optical networking has undergone tremendous changes and the trend clearly shows an evolution path towards lower cost (CAPEX and OPEX) and higher capacity. Apart from these costs, there are concerns regarding the physical space requirements, energy consumption and heat dissipation. These changes have been governed by developments of networking capabilities (e.g. more wavelengths, higher line rates) and emerging applications (e.g. video services). The optical network evolution was focused on providing more capacity in a cost-effective manner. With respect to the optical transmission systems, this evolution can be translated to denser WDM transmission systems (i.e. 80–160 wavelengths per fiber) operating at higher line rates (e.g. 10 Gpbs, 40 Gbps or even 100 Gbps), and coarser granularities at switching level [1]. Furthermore, providing static and high-capacity pipes is no longer sufficient to address the demands of emerging dynamic applications. Therefore a dynamic and configurable optical layer and control plane, which is able to serve dynamic requests, is the direct consequence of the mentioned trend. However, all of these requirements should be addressed utilizing a cost-effective solution.

Optical network architectures are evolving from traditional opaque networks toward all-optical (i.e. transparent) networks as depicted in Fig. 1. In opaque networks, the optical signal carrying traffic undergoes an optical–electronic–optical (OEO) conversion at every switching or routing node. Given the practical and economical considerations, the transmission reach of optical signals is limited (e.g. 2000–2500 km) [2]. To go beyond this transparent reach of optics limit, signal regeneration is essential to
re-amplify, re-shape and re-time the optical signal (also known as 3R). Regeneration simply improves the quality of the optical signal. The OEO conversion enables the optical signal to reach long distances; however, this process is quite expensive due to several factors such as the number of regenerators required in the network, the dependency of the conversion process to the connection line rate and also to the modulation format. An OEO node, especially one based on electronics, will have its own scalability issues related to cost, space requirements, power consumption and heat dissipation. As the size of opaque networks increases, network designers and architects have to consider more electronic terminating and switching equipments, which presents challenges in cost, heat dissipation, power consumption, required physical space, and operation and maintenance costs.

One approach to address these issues is the use of sparsely placed electrical or optical regenerators [3]. In principle, regeneration can be accomplished completely in the optical domain (e.g. [4,5]); however, regeneration in the electronic domain (i.e. OEO conversion) is still the most economic and reliable technique. All-optical regeneration is a relatively new technology that is not mature enough and is still an area of active research on many fronts. The lack of practical all-optical regeneration, gives rise to the intermediate optical network architectures, which are identified as translucent [6] or optical-bypass [7] networks. Translucent network architectures have been proposed as a compromise between opaque and all-optical networks. In this approach, a set of sparsely but strategically placed regenerators is used to maintain the acceptable level of signal quality from the source to its destination. This approach in fact eliminates much of the required electronic processing and allows a signal to remain in the optical domain for much of its path. Since optical technology can operate on a spectrum of wavelengths at once and also can operate on wavelengths independent of their line rate, keeping the signals in optical domain brings a significant cost reduction due to removal of electronic processing equipments [8]. This removal also paves the way for lower power consumption, heat dissipation and site space requirements. Optical-bypass core WDM networks using reconfigurable optical add/drop multiplexers (ROADMs) and tunable lasers appear to be on the road towards widespread deployment and could evolve to all-optical mesh networks based on optical cross connects (OXCs) in the coming future. No matter which architecture is considered, the main goal of these architectures is to provide the required infrastructure for end-to-end connection establishment.

In optical networks, a lightpath is an optical path established between a pair of source-destination nodes. The demand set (or traffic matrix) in the network is the collection of lightpaths that must be established. The term “demand” represents an individual request for lightpath establishment. In the context of network planning, some demands are permanent and are referred to as permanent lightpath demands (PLD) or Static demands. The other categories of demands are defined as dynamic lightpath demand (DLD) or dynamic demands for short, in which demand requests have a finite lifetime (i.e. start and end) [9]. In this case two variants of DLD can be distinguished (see Fig. 2):

- scheduled lightpath demands (SLD): the activation time (date) and lifetime of this demands are known in advance. Provisioning of layer 1 virtual private networks (VPNs) falls under this category. Since SLDs are pre-planned, they may be considered as a whole during the network planning or operation phase.
Ad hoc lightpath demands (ALD): this category of demands is characterized by the fact that their arrival time (date) and also their lifetime are not known a priori. These two parameters (i.e. arrival time and duration) may be modeled in general by two random processes.

Each lightpath is created by allocating a single wavelength (assuming wavelength conversion is not present) throughout the path. If the allocated wavelength for a given lightpath remains the same across all fiber links that it traverses, then the routing and wavelength assignment (RWA) [10,11] is said to satisfy the wavelength continuity constraint. However, if a switching node is equipped with a wavelength conversion facility, then the wavelength continuity constraint disappears and the routing problem will be reduced to normal routing in circuit-switched networks, where the only limiting factor is the number of available wavelengths on each link.

Previous studies have already investigated the RWA problem as summarized in [10]. The RWA problem is known to be NP-Complete. In much of these works, the assumption is that the network is truly all-optical, where all intermediate regenerations (i.e. OEO conversion) are eliminated. Also in most RWA proposals the optical layer is considered as a perfect transmission medium and therefore all outcomes of the RWA algorithms are considered valid and feasible. The reality is that the actual performance of the system may be unacceptable for some of the lightpaths. For this reason the incorporation of physical layer impairments in transparent/switched optical network planning and operations has recently received some attention from research communities. The physical layer impairments are either considered as constraints for the RWA decisions (i.e. physical layer impairment constrained (PLIC-RWA)) or the RWA decisions are made considering these impairments (i.e. physical layer impairment aware (PLIA-RWA)). In the latter, it is possible to find alternate routes considering the impairments, while in the former the routing decisions are constrained by physical layer impairments. However, for simplicity, in the rest of this paper we use the generic PLI-RWA term. Since the reach of optical signals is limited, some amount of intermediate regeneration is necessary in carrier backbone networks. Therefore the PLI-RWA problem will be inevitably coupled with regeneration placement problem [13], in which the network designers are trying to plan and design translucent (or optical-bypass) networks with optimal number of regeneration sites for a given network topology and demand set (i.e. traffic matrix). The regenerator placement problem is also known to be NP-complete [13,24].

In addition to analytical and simulation techniques for modeling physical layer impairments, monitoring techniques are required for measurements, which potentially can enhance the PLI-RWA algorithms. The monitoring could be implemented on the impairment level (optical impairment monitoring – OIM) or at the aggregate level where the overall performance is monitored (optical performance monitoring – OPM) [14,15].

The main motivation behind this work is the lack of a comprehensive literature survey on PLI-RWA algorithms. In order to report the state-of-the-art and current proposals, we collected more than a hundred recent papers, and considering 28 different metrics we reviewed and ranked the related papers. The result of our study is compiled and reported in this work, which is to the best of our knowledge the only comprehensive literature survey for studies in the area of PLI-RWA.

The remainder of this paper is organized as follows. In Section 2, we examine various physical layer impairments in optical networks and based on our survey we present a categorization of these impairments. Section 3, of our survey reports on various PLI-RWA algorithms and we also provide different classification of approaches that have been proposed for this problem. Possible extensions to the control plane of the optical networks are presented in Section 4. Discussions and some remarks are compiled in Section 5 and conclusions are given in the last section of this paper.

2. Physical layer impairments

As optical signals traverse the optical fiber links and also propagate through passive and/or active optical components, they encounter many impairments that affect the signal intensity level, as well as its temporal, spectral and polarization properties. Physical layer impairments can be classified into linear and nonlinear effects. Linear impairments are independent of the signal power and affect each of the wavelengths (optical channels)
individually, whereas nonlinear impairments affect not only each optical channel individually but they also cause disturbance and interference between them [16,17].

2.1. Linear impairments

The important linear impairments are: fiber attenuation, component insertion loss, amplifier spontaneous emission (ASE) noise, chromatic dispersion (CD) (or group velocity dispersion (GVD)), polarization mode dispersion (PMD), polarization dependent losses (PDL), crosstalk (XT) (both inter- and intra-channel), and filter concatenation (FC).

Optical amplification in the form of EDFAs always degrades the optical signal to noise ratio (OSNR). The amplifier noise is quantified by noise figure (NF) value, which is the ratio of the optical signal to noise ratio (OSNR) before the amplification to the same ratio after the amplification and is expressed in dB [16].

Chromatic dispersion causes pulse broadening, which affects the receiver performance by: (1) reducing the pulse energy within the bit slot and (2) spreading the pulse energy beyond the allocated bit slot leading to inter-symbol interference (ISI). CD can be adequately (but not optimally) compensated for on a per link, and/or at transmission line design time [16,18–20].

PMD is not an issue for most type of fibers at 10 Gbps, however it become an issue at 40 Gbps or higher rates [17,20–23]. In general, in combination with PMD there is also polarization dependent loss (PDL). It can cause optical power variation, waveform distortion and signal-to-noise ratio fading.

Imperfect optical components (e.g. filters, demultiplexers, and switched) inevitably introduce some signal leakage either as inter-channel [16,20] (also incoherent [20] or out-of-band [24]) or intra-channel [16,20] (or intraband [24]) crosstalk in WDM transmission systems.

Filter concatenation is the last physical impairment that we consider and define in this category. As more and more filtering components are concatenated along the lightpath, the effective pass band of the filters becomes narrower [20]. This concatenation also makes the transmission system susceptible to filter passband misalignment due to device imperfections, temperature variations and aging.

2.2. Non-linear impairments

Important non-linear impairments can be summarized as self phase modulation (SPM), cross phase modulation (XPM), four wave mixing (FWM) [26], stimulated brillouin scattering (SBS), and stimulated Raman scattering (SRS) [25].

The nonlinear phase shift manifests as phase modulation. In SPM the phase of the signal is modulated by its own intensity; while in XPM the signal phase is modulated by the intensity of other signals [16]. The primary effect of these impairments is pulse broadening in frequency domain without changing the shape of the signal.

SBS and SRS involve non-elastic scattering mechanism [16,25]. These impairments set an upper limit on the amount of optical power that can be launched into an optical link.

2.3. Classification of physical impairments

Fig. 3 depicts the classification of physical layer impairments, considering linear and non-linear categories. We can also classify the physical layer impairments to be static or dynamic impairments, considering the dependence of impairments behavior on external factors such as aging, temperature, and physical stress.

In order to incorporate the physical layer impairments effects in the RWA algorithms two general models have

![Fig. 3. Classification of physical layer impairments.](image-url)
been reported. These approaches are: (1) analytical models, (2) Hybrid (analytical models accompanied by simulation results or optical impairment monitoring techniques). In the former the physical layer impairments are evaluated using closed-form formula and in the latter some simulation results or real-time impairment monitoring are also considered for the evaluation of the physical layer performance. Among a number of measurable optical transmission quality attributes (e.g., optical power, OSNR, CD, PMD, Q factor) [13,42,44] shows the best suitability as an integrated metric for routing algorithm, due to its close correlation with BER. Q factor is sensitive to all forms of BER impacting impairments. It is also possible to evaluate the quality of transmission using hybrid or experimental-based models. Authors of [29] report the quality of transmission (QoT) function. This function is obtained by considering experimental measurements. Authors of [9] have also used a similar function. Table 1 summarizes the reported performance evaluation techniques along with the relevant references.

Some physical impairments have strong dependency on bit rates, modulation formats and type of amplifier that are considered as part of the optical links model. Table 2 summarizes different modulation formats, type of optical amplifiers and also bit rates that are assumed and reported in surveyed studies.

As indicated in Table 2, there are few works in the state-of-the-art surveyed papers that consider the advanced modulation formats and also Raman amplifiers in their studies. Also higher bit rates and the challenges that they will introduce are areas that require more research.

### 3. PLI-RWA algorithms

In this section we present the algorithmic approach for solving the PLI-RWA problem and the classification of the different PLI-RWA proposals in the surveyed literature. This section also includes a description of the performance metrics and the methods adopted to evaluate the different proposals.

### 3.1. Algorithmic approach

In general the algorithmic approach for the PLI-RWA problem can be categorized either as sequential approach based on some heuristic or meta-heuristic algorithms, which usually give a sub-optimal solution, or combinatorial approach, which searches for an optimal solution.

The classic RWA problem (i.e., without PLI constraints) is NP-complete [10] and thus its optimal solution cannot be found in polynomial time using any known algorithm. The PLI-RWA problem introduces additional difficulty to the RWA problem since it involves a number of physical layer-related constraints. To alleviate these obstacles the RWA problem can be decomposed into two sub-problems, namely, a routing (R) problem, i.e. choice of a suitable path, and a wavelength assignment (WA) problem, i.e. allocation of an available wavelength for the selected path. Then each problem can be solved separately. When treating routing and wavelength assignment steps separately and individually, each step can be further broken into two components: (1) search and (2) selection. The first step concerns the search for a set of candidate paths/wavelengths, which may be also the subject of appropriate ordering consideration. The second step is a decision function that operates on the given candidate set.

As already mentioned, further simplification of the PLI-RWA problem can be achieved with the application of a heuristic (or meta-heuristic) algorithm, which can be used to solve any of the R and/or WA sub-problems. Although such a decomposition of the RWA problem does not guarantee its optimal solution, still the computation time can be reduced considerably.

#### 3.1.1. Heuristics

Since the PLI-RWA problem involves additional physical-layer constraints, which in most cases are verified using complex analytical physical models of physical impairments, the complexity of the proposed algorithms

<table>
<thead>
<tr>
<th>Performance evaluation technique</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analytical models</strong></td>
<td></td>
</tr>
<tr>
<td>Linear impairments</td>
<td>[17,19,23,27,28,30,33–38,43,46,49,51–54,56,57,59,66,72–75]</td>
</tr>
<tr>
<td>Non-linear (and linear) impairments</td>
<td>[26,31,32,41,47,48,50,60]</td>
</tr>
<tr>
<td><strong>Hybrid approach</strong></td>
<td></td>
</tr>
<tr>
<td>Analytical and simulation (linear and/or non-linear impairments)</td>
<td>[13,18,24,35,40,42,44]</td>
</tr>
<tr>
<td>Analytical and Monitoring/Experiments (linear and/or non-linear impairments)</td>
<td>[9,29,73]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2: Considered modulation format, amplifiers, and bit rates.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Modulation</strong></td>
</tr>
<tr>
<td>OOK (NRZ/RZ)</td>
</tr>
<tr>
<td>[13,19,24,29,41,47–49,53,60]</td>
</tr>
</tbody>
</table>
are significantly important. Therefore most of the PLI-RWA algorithms, reported in the literature are based on simple heuristics.

3.1.1.1. Routing sub-problem. Regarding the routing subproblem, a variety of heuristics algorithms can be found in the literature and, in general, they are based on the shortest path (SP) routing algorithm (e.g., Dijkstra algorithm). Among these algorithms, two classes of algorithms can be distinguished, namely, single-path routing algorithms and multi-path routing algorithms. The latter are also known as \( k \)-shortest path (\( k \)-SP) algorithms. In any SP-based algorithm there is a cost (or weight) parameter assigned to each network link. This parameter should have additive properties and it is used by the algorithm to find a path of a minimum overall cost. The link cost may simply represent a hop distance (i.e., unique and similar cost for all links) or it may correspond to the link state and then it involves link congestion or physical impairment-related information.

Regarding single-path routing, a number of PLI-RWA proposals follow the minimum hop SP approach as reported in [13,17,31,34,35,40,46,48,49]. As for the calculation of PLI-aware link costs, the simplest metric which has additive properties is the physical distance [46]. Other PLI-RWA algorithms use the link cost that is a function of the residual dispersion parameter [35], the FWM crosstalk [26], the Q-factor [46], or the noise variance [47]. In this last case, both linear and non-linear impairments are represented by the noise variance. This solution may be preferable since noise variances are additive hop by hop.

Some PLI-RWA methods introduce modifications to the SP routing algorithm. For instance, the minimum coincidence and distance (MINCOD) algorithm [33] tries to compute the paths that minimize the distance and the number of shared links, while the least-congested algorithm [40] aims at balancing the network load. Besides, an impairment-constraint-based SP algorithm, which takes into account the utilization of network resources and the physical impairment due to FWM crosstalk was proposed in [26]. Finally, an adapted Bellman–Ford shortest path algorithm that deals with multi-objective, constraint-based lightpath provisioning was proposed in [18].

PLI-RWA algorithms based on multi-path routing algorithms operate on a set of pre-calculated alternative paths. Since these paths are in usual the shortest paths, such class of multi-path algorithms can be also referred to as \( k \)-shortest path (\( k \)-SP) algorithms. In some cases the set of candidate paths is restricted to disjoint paths [24].

Similar to the single-path routing, in the multi-path routing the cost metric can be related to the hop count as utilized in [9,30,43,50,52] or can be PLI-aware. The distance metric is the simplest PLI-aware link cost [24,59]. Also, other physical impairments, such as PMD, ASE noise, XT, CD, and FC, can be adopted to represent the link cost [53]. Besides, many proposals consider a Q-factor as a link cost. This cost can be based on real-time Q-factor measurements collected from devices [39] or can be calculated analytically either as the worst case Q-factor penalties [24] or taking into account linear [42] and non-linear impairments [44]. More complex link cost formulations may combine a number of parameters, such as information about regenerating modules, the number of available and total wavelengths, and the link length [37].

Once candidate paths are found, the selection of an appropriate path is performed either sequentially or in parallel. In the former a sequence of re-attempts is performed until the first available path that complies with the given performance requirements is found [43,52,70]. In the latter the path which is the most suitable according to a given decision criteria is selected [30,54]. The search among multiple alternative paths can be implemented by the breath-first search (BFS) algorithm [43], and [54]. BFS tries to examine all nodes of a network graph in some systematic way in order to explore all possible solutions.

3.1.1.2. Wavelength assignment sub-problem. The wavelength assignment subroutine operates on a set of candidate wavelengths that are given on a previously selected routing path (or paths). The set may be ordered, according to a given policy, or unordered, i.e., the wavelengths are treated in a round-robin way. Wavelength ordering was proposed in [55] as a technique to select the wavelength with lower number of adjacent-port crosstalk terms. As an extension to this method, the WA algorithm in [48] initially considers the wavelengths that are most separated in terms of frequency. Then wavelengths are analyzed in an optimal order to maximize the frequency separation.

Given a set of candidate paths, the wavelength selection phase can be performed either sequentially or in parallel. This is similar to the routing sub-routine. In the sequential approach, the first non-occupied wavelength that satisfies given network-layer and physical-layer constraints is selected. Such a first-fit (FF) selection method has been considered in a large number of PLI-RWA proposals [9,13,18,31,35–37,40,43,47,51,56,57,59]. On the contrary, some PLI-RWA algorithms try to look through all of the candidate wavelengths so as to find the best-fit (BF), i.e., the most appropriate one [18,34]. For instance, the wavelength of the lowest utilization in the network can be selected based on the network state information given at the sources node, as in the least-loaded algorithm [33]. Finally, a random selection, which means choosing randomly amongst the available wavelengths, can be performed [17,47]. It is well known that wavelength blocking probability of a random WA algorithm is worse than that of the FF algorithm [10]. Nonetheless, since the random algorithm tends to geographically spread the wavelength use across the network, the crosstalk effects might be limited [17].

Some of the proposed WA algorithms make a decision based strictly on physical layer impairments. An example can be a PLI-aware algorithm presented in [26] which aims at minimizing the FWM crosstalk effect. This algorithm has been proposed in two versions, namely, to perform either FF or BF wavelength selection. Another two algorithms can be found in [60] and while one of them focuses on the selection of the lightpath with the highest Q-factor, the other addresses the fairness issue and it also tries to minimize the impact of this lightpath on the already established lightpaths.
3.1.1.3. Routing and wavelength assignment. Apart from separate R and WA solutions, there are some heuristics that intend to solve the PLI-RWA problem jointly. To achieve this goal the A* (A star) algorithm, which is a shortest path algorithm derived from the Dijkstra algorithm, has been proposed in [29]. The A* algorithm relies on a layered network graph that is derived from a network graph by multiplication of links and vertices by the number of corresponding wavelengths. Thanks to the layered representation of links and wavelengths in a single graph the algorithm is able to find an appropriate lightpath in one algorithmic step.

Another example can be the minimum crosstalk (MC) algorithm [38]. For each wavelength, MC runs a SP algorithm to find candidate routes. The link weights are constant and equal to the physical link lengths. For each candidate route, the number of crosstalk components along the route is calculated. Among all the candidate routes, it chooses the route at the wavelength with the minimum crosstalk intensity.

Finally, the Best-OSNR algorithm that jointly assigns to a given request a path and a wavelength in order to maximize the OSNR, was proposed in [40]. In Table 3 we summarize the reported heuristic algorithms.

<table>
<thead>
<tr>
<th>(Sub-)Problem</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Routing</strong></td>
<td></td>
</tr>
<tr>
<td>Single-path</td>
<td></td>
</tr>
<tr>
<td>Hop-based shortest path</td>
<td>[13,17,31,34,35,40,46,48,49]</td>
</tr>
<tr>
<td>PI-aware shortest path</td>
<td>[26,35,46,47]</td>
</tr>
<tr>
<td>Modified shortest path</td>
<td>[18,26,29,33,40]</td>
</tr>
<tr>
<td>Multi-path (route calculation)</td>
<td></td>
</tr>
<tr>
<td>Hop-based k-SP</td>
<td>[9,30,43,50,52]</td>
</tr>
<tr>
<td>PI-aware k-SP</td>
<td>[24,37,42,44,53,59]</td>
</tr>
<tr>
<td>Multi-path (route selection)</td>
<td></td>
</tr>
<tr>
<td>Sequential (re-attempt)</td>
<td>[43,52,70]</td>
</tr>
<tr>
<td>Parallel (best one)</td>
<td>[30,54]</td>
</tr>
<tr>
<td><strong>Wavelength assignment</strong></td>
<td></td>
</tr>
<tr>
<td>First-fit</td>
<td>[9,13,18,31,34–37,39,40,43,47,51,56,57,59]</td>
</tr>
<tr>
<td>Best-fit</td>
<td>[18,34]</td>
</tr>
<tr>
<td>Least-loaded</td>
<td>[33]</td>
</tr>
<tr>
<td>Random</td>
<td>[17,47]</td>
</tr>
<tr>
<td>PI-aware</td>
<td>[26,41,59,60]</td>
</tr>
<tr>
<td><strong>Routing and wavelength assignment</strong></td>
<td></td>
</tr>
<tr>
<td>Minimum crosstalk</td>
<td>[38]</td>
</tr>
<tr>
<td>Best-OSNR</td>
<td>[40]</td>
</tr>
</tbody>
</table>

3.1.2. Meta-heuristics

Apart from the heuristic-based algorithms, there is a class of PLI-RWA algorithms that exploit meta-heuristic methods. Meta-heuristics are very attractive as far as they do not involve complex mathematical formulations and, at the same time, they allow the convergence to an optimum solution through successive iterations.

The ant colony optimization (ACO) is one of meta-heuristics applied to solve the PLI-RWA problem [58] [57,59]. ACO is characterized by ant-like mobile agents that cooperate and stochastically explore a network. The agents build iteratively solutions based on their own information and on the traces (called pheromones) left by other agents in network nodes. In the proposed PLI-RWA method, the ACO algorithm calculates the path on a hop-by-hop basis. The next hop is calculated based on pheromone values of the node, which accounts for the OPM of the links. The algorithm is capable of the distributed calculation of a multi-constrained path under restrictions resulting from ASE noise and optical power budget.

Another PLI-RWA algorithm [61] makes use of a genetic algorithm (GA). A GA operates on a set of solutions called population. In each iteration appropriately selected solutions from one population are used to form, through a number of operations, a new population that is expected to be a better one. The proposed PLI-RWA algorithm attempts to compute a lightpath in such way that the average blocking probability and the usage of optical devices, such as wavelength converters and amplifiers, is minimized. Both ASE noise and PMD are considered as physical impairments.

To solve the problem of survivable lightpath provisioning in a translucent network the Tabu-Search (TS) meta-heuristic has been applied [37]. TS is a neighborhood search method which tries to avoid local minimum by accepting worse solutions and by using the solutions’ search history. In the proposed solution, the TS algorithm operates on a set of k-SP, where k is dynamically changing according to the direction of improvement. This adaptive feature improves the efficiency of the method. As for the physical constraints, both PMD and ASE noise are used.

Finally, in [32] the authors propose a predictive algorithm (PA) for the PLI-RWA problem. The main idea of this algorithm is to apply the branch prediction concept originally used in the computer architecture area. In optical networks, the algorithm selects the lightpath based on the history of previous connection requests. The main advantage of this algorithm is that it can be used in distributed routing and it does not need any update messages with global network information in order to compute the lightpath. The physical impairment considered by this algorithm is the maximum transmission distance. In Table 4 we summarize the reported meta-heuristic algorithms.

### Table 4

<table>
<thead>
<tr>
<th>Meta-heuristics</th>
<th>Optimization methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACO</td>
<td>GA</td>
</tr>
<tr>
<td></td>
<td>TS</td>
</tr>
<tr>
<td></td>
<td>PA</td>
</tr>
<tr>
<td></td>
<td>M(ILP)</td>
</tr>
<tr>
<td></td>
<td>LP</td>
</tr>
<tr>
<td></td>
<td>DP</td>
</tr>
</tbody>
</table>

3.1.3. Optimization methods

The last class of methods considered for PLI-RWA is based on the network optimization theory. The network optimization methods are usually appropriate for off-line optimization of network resources as well as for on-line and centralized lightpath provisioning. Among the solutions presented in the literature, most of them have been proposed for transparent networks.

In [53] a link-path formulation to solve an integer linear programming (ILP) problem of RWA in a transparent network is proposed. A set of k paths is pre-calculated with
the assistance of a SP algorithm, which uses either a single physical impairment [53] or a Q-Penalty [42] as the link cost parameter. Presented ILP formulation takes into account the existence of sparse wavelength-conversion capable nodes in the network.

Some more specific problems involving PLI constraints into the optimization problem were studied as well. A mixed-ILP (MILP) formulation for the RWA problem of multicast connections, while considering optical power constraints, is proposed in [51]. Authors in [58] consider algorithms for the logical topology design and traffic grooming problem in WDM networks with router interface constraints as well as optical constraints. Their approach is based on a linear program which is NP-complete. They also introduce heuristic algorithms which use a graphical modeling tool. Also, an ILP formulation for the problem of traffic grooming in optical virtual private networks with the BER constraint is presented in [62].

In the case of translucent networks, the problem of regenerator placement with constraints on OSNR is solved using an ILP formulation [63]. A solution to similar problem, considering BER constraints is reported by applying dynamic programming (DP) technique [64]. Moreover, the problem of survivability in lightpath provisioning in a translucent network is solved in [37] using an ILP formulation. In this work PMD and ASE noise are considered.

In [70] the implementation of an LP solver in a path computation element (PCE) is reported. The implemented objective function minimizes the maximum link bandwidth utilization. As a result the routes which satisfy the required constraint in terms of bandwidth and optical signal quality can be found. In Table 4 we summarize the reported optimization algorithms.

### 3.2. PLI-RWA proposals

When selecting a lightpath (route and wavelength), a PLI-RWA algorithm for a transparent or translucent network has to take into account the physical layer impairments as well as the wavelength availability. With static traffic, the entire set of connection requests is known in advance, and the static (off-line) RWA problem of setting up these connection requests is named the permanent lightpath establishment (PLD) problem. In a dynamic traffic scenario the connections are requested in some random fashion, and the lightpaths have to be set up as needed (introduced as SLD or ALD in Section 1 above). In static (off-line) case, there is enough time between the planning and provisioning processes such that any additional equipment required by the plan can be deployed. In this context the main goal is to accommodate the whole demand set. In dynamic traffic scenario, there is little time between planning and provisioning, and demands are generally processed one at a time. It is assumed that the demand set must be accommodated using whatever equipment is already deployed in the network. Thus, the PLI-RWA proposal must take into account any constraints posed by the current state of the network, which may force a demand to be routed over a sub-optimal path. One big challenge for PLI-RWA algorithms is the QoT-awareness, in the sense that they must ensure (during admission control) that all lightpaths in the network meet a QoT (e.g. BER) constraint without disrupting previously established lightpath.

The effect of the existing connections in the PLI-RWA decision is rarely taken into account in the proposed algorithms in the literature. Some works address this problem considering the crosstalk due to the already established connections [38,41,48]. For example in [41], the HQ (Higher Q) and MMQ (maximize minimum) algorithms try to minimize the effect of new crosstalk when establishing a lightpath. The MC (minimum crosstalk) [38] wavelength assignment selects the wavelength with minimum crosstalk intensity due to the already established connections. A different approach is considered in [47], where, in the lightpath selection, the BER of the selected and affected lightpaths are taken into account in the lightpath establishment. Few works address the problem of selecting the lightpath considering the effect of selected lightpath on the possible future demands. In [19,35], the dispersion optimised impairment constraint-based (DOIC) PLI-RWA algorithm assigns the wavelength with the lowest residual dispersion. This is done to increase the wavelength availability for the upcoming demands.

The re-routing feature is even rarer than the effect of the existing connections on upcoming demands, mainly due to its complexity. Re-routing refers to the re-computation and re-establishment of already established connections when a new lightpath is established. In [73], re-routing is utilized to perform the restoration of label switched paths (LSPs) channels based on a threshold of the OSNR values.

There are several heuristic algorithms proposed in the literature dealing with the wavelength assignment sub-problem, such as random, first-fit (FF), least-used (LU), etc. [10].

When the PLIs are introduced in the RWA algorithms, three main approaches have been considered in the recent literature: (a) compute the route and the wavelength in the traditional way and finally verify the selected lightpath considering the physical layer impairments; (b) considering the PLI values in the routing and/or wavelength assignment decisions and (c) considering the PLI values in the route and/or wavelength assignment decision and finally also verify the quality of the candidate lightpath. These cases and their various combinations are depicted in Fig. 4.

In case A-1 the route and the wavelength are selected without considering the PLI constraints, but after the verification phase, the wavelength assignment decision can be modified. In case A-2 the route is computed without taking into account the physical impairments, but there are re-attempts of computing the route if the PLI constraints are not met for the candidate route(s); finally there is a final phase of traditional wavelength assignment. In case A-3 the route and the wavelength are computed using classical RWA algorithms that are unaware of the PLI constraints (selecting both route and wavelength in one step or selecting first the route and then the wavelength) and there is a final step of checking the PLI constraints in order to possibly change the RWA decision.

Different works (e.g. [13,27,30]) have followed the A-2 approach. In [13], a combination of PLI-RWA algorithm
and regenerator placement for dynamic traffic is proposed. This algorithm computes paths between any combinations of nodes under the constraint of a minimal $Q$ (quality) value. Authors in [27] propose a modification of the Bellman–Ford algorithm to compute the path with the minimum number of hops with a certain (monetary) cost limit. At destination the physical layer requirements with different levels of agreement are checked. Also for static traffic, the proposed algorithm in [30] computes $k$-shortest paths considering the costs of the links associated with impairments and OEO devices; the final path is chosen in a way that satisfies a minimum $Q$ value. The path establishment process is also taken into account in [30] and two methods (sequential and parallel) are compared. In the sequential method, the source node sends out a single PATH message containing optical properties which are checked at destination; if the source node receives back a path-error message it will try another path from the list of candidate paths. In the parallel method, the source node sends out $k$ PATH messages and the destination node makes the path selection.

The A-3 approach is utilized in [9,23,29,32–34]. The LERP (lightpath establishment and regenerator placement) algorithm is proposed in [9] as a PLI-RWA and regenerator placement algorithm. It computes $k$ alternate paths, then the wavelength is assigned using the FF or random method; and finally there is a phase of testing the $Q$ values of the lightpath and placing regenerators if any is needed. In [23] the proposed PLI-RWA algorithm is simply the SP combined with FF wavelength assignment, then the quality of the candidate lightpath is verified in terms of maximum BER. In [29], the $A^*$ algorithm is employed for RWA; it computes $k$-SPs and the path is chosen satisfying a minimum $Q$ value and having the smallest cost in terms of fiber and regenerator utilization. A variation of A-3 approach is found in [31], in this work the wavelength of the path is initially assigned by means of FF algorithm and then the SP for the selected wavelength is computed. Afterwards, the final phase verifies the level of OSNR and the pulse broadening. In [32] the RWA is performed based on the SP algorithm for each available wavelength and then the lightpath with the highest $Q$ value is selected. In [33], the lightpath selection is done using a predictive approach and taking into account the possible inaccuracy in the wavelength availability information; then a verification of the maximum transmission distance achievable for that wavelength is performed. Finally in [34], two phases compute the lightpath. First a lightpath computation step (best path selects the SP among all the available wavelengths or FF selects the SP for the first available wavelength) and then a lightpath verification step based on a BER threshold is performed. In case A-1 the physical impairments are only verified in wavelength assignment process and the possibility of selecting other routes to get better QoT is ignored. In contrary in case A-2 different routes are considered to meet required QoT, while eventually the wavelength is assigned without considering the impact of physical impairments. The last case (i.e. A-3) does not verify the impact of physical impairments neither in routing nor in wavelength assignment process. Only the QoT is verified after finding a potential solution and if the answer is not satisfactory the whole process is repeated.

In general the approaches in group B (cases B-1, B-2, and B-3) address the RWA problem considering the physical layer information: in case B-1 the route is computed using PLI constraints; in case B-2 these constraints are considered in the wavelength assignment process; finally in case B-3 the PLI constraints are taken into account in both, route and wavelength selection. Some of the works that present this approach use the physical layer information as weight of the links, in order to compute the minimum cost lightpath.

Fig. 4. Various PLI-RWA approaches.
The minimum crosstalk (MC) wavelength assignment in [38] follows the B-2 approach. The MC selects the wavelength with minimum crosstalk intensity.

The B-3 variations, where the PLI constraints are taken into account in both route and wavelength selection process, is found in [19,39–41]. The DOIC algorithm presented in [19] considers the PLI constraints related to CD and OSNR in the route decision process, and also the residual dispersion range in the wavelength assignment. In [39] the shortest cost path or the k link disjoint shortest cost paths are computed considering the Q value as the link cost, which is obtained from real-time Q measurements. Depending on the network conditions the final decision is taken according to the wavelength balancing efficiency or according to the Q factor value. The Best OSNR RWA algorithm proposed in [40] selects the lightpath with maximum OSNR. In [41] two PLI-RWA algorithms are proposed, the HQ (highest Q) and the MMQ (maximum minimum Q) algorithms. The HQ selects the lightpath with the highest Q value and MMQ selects the lightpath that maximizes the minimum Q value among the paths that are affected by the establishment of the new lightpath. Note that MMQ takes into account the impact of the lightpath decision on the already established connections (in section 5 the related literature related to this impact is presented). In general different variations of case B (i.e. B-1, B-2, and B-3) consider the impact of physical impairments in routing and/or wavelength assignment process, however this scheme [41] does not try to verify the QoT of the solution or find the optimum one.

Case C is a combination of the two previous approaches. The PLI constraints are taken into account in the routing (case C-1), or in the wavelength assignment (case C-2) or in both (case C-3); but there is a final phase of verification of the PLI constraints that enables the re-attempt process in the lightpath selection phase.

Examples of approach C-1 are found in [24,30,42–44,46,47]. Two algorithms proposed in [24] compute k-shortest paths (k = 3) using as link costs the worst case of Q values. Then, the whole Q value is verified at destination comparing it with a threshold. The Q threshold is an offline value in case of static routing, or a dynamic Q value, based on current status of the network, for dynamic routing. The third proposed algorithm in [30] selects the route between source and destination in a hop by hop manner. In each node the physical feasibility is verified; furthermore some PLI constraints are also checked at the destination. In [42] k-shortest paths are computed considering the network and physical characteristics such as the link costs; finally there is a validation of the PLI constraints considering the Q factor, among the previously computed paths. A similar approach is followed in [43], where alternate paths are selected by pruning in the topology those links that are not fulfilling the dispersion and ASE constraints. The lightpath selected is the one with least number of links. Moreover other PLI constraints are checked at destination. In [44] the Q-factor penalty is used as the link cost in the network, in order to compute the k-shortest paths. Finally the BER is computed and verified at destination; the selected path is the one with lowest BER value considering a BER threshold. The proposed PLI-RWA algorithms in [46] utilizes link costs related to the Q factor (as 1/Q or 1/Q^2) for computing the SP; then a PLI constraints verification phase considers the availability of the lightpath or otherwise the maximum reachable node. In [47] authors propose a variation of sub-case C-1, in which the wavelength is initially selected by means of FF (unaware of PLI constraints) algorithm, and then SP for that wavelength is computed considering the noise variance of the PLI as the link cost. Finally, lightpaths which cause a BER value higher than a given threshold for the new lightpath or for other already established lightpaths, are discarded.

In [48] authors propose the C-2 approach, where the SP is computed unaware of PLI constraints and the wavelength assignment uses a wavelength order considering the PLI constraints. Finally, there is a verification of the quality of the lightpath in terms of minimum Q value. The proposal in [49] is very similar to the mentioned approach, however, a BER threshold is considered at the destination to validate the quality of the lightpath.

An example of sub-case C-3, where the PLI constraints are taken into account in both routing and wavelength assignment processes, is found in [50]. In this work, k-shortest routes are computed considering a Q-Penalty value as the link costs, and the selected wavelength is the one that maximizes the Q value; and finally the quality of transmission (QoT) is verified.

Different schemes in Case C (i.e. C-1, C-2, and C-3) not only try to consider the impact of physical impairments, but also verify the QoT of the solutions and also try to find optimum solution. Obviously the cost of this scheme is its complexity (in terms of running time). Table 5 provides a summary of these cases and the related surveyed papers.

After this summary of recent PLI-RWA algorithms, we have to mention that few works ([19,38,48,50]) utilize or propose specific wavelength (WA) assignment algorithms taking into account the PLI constraints. Different combina-

Table 5
Summary of PLI-RWA proposals.

<table>
<thead>
<tr>
<th>Case</th>
<th>Indicative references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case A-1</td>
<td>N/A</td>
</tr>
<tr>
<td>Case A-2</td>
<td>[13,27,30]</td>
</tr>
<tr>
<td>Case A-3</td>
<td>[9,23,29,32–34]</td>
</tr>
<tr>
<td>Case B-1</td>
<td>[35–37]</td>
</tr>
<tr>
<td>Case B-2</td>
<td>[38]</td>
</tr>
<tr>
<td>Case B-3</td>
<td>[19,39–41]</td>
</tr>
<tr>
<td>Case C-1</td>
<td>[24,30,42–44,46,47]</td>
</tr>
<tr>
<td>Case C-2</td>
<td>[48,49]</td>
</tr>
<tr>
<td>Case C-3</td>
<td>[50]</td>
</tr>
</tbody>
</table>
tion of the sub-cases presented in Fig. 4 may also be found in the literature.

### 3.3. Wavelength conversion

Apart from the physical-layer constraints usually there is a wavelength-continuity constraint imposed on the RWA problem in optical networks. This constraint means that a given lightpath connection should be composed of identical wavelengths on the links traversed by the lightpath. Such requirement may affect both the network performance and the complexity of RWA algorithm since the setup of a new lightpath is conditioned on the availability of the same wavelength in a number of links. The wavelength-continuity constraint can be relaxed in the nodes that are capable of wavelength conversion, thus improving the connection blocking probability. In practice the wavelength conversion can be realized in switching nodes either by means of a dedicated all-optical device or with the assistance of an optical–electrical–optical (OEO) signal regenerator. The OEO regenerator converts an input wavelength to an electronic signal and then converts it back onto another wavelength. Because all-optical wavelength converters are still immature and very expensive, the OEO wavelength conversion becomes a viable alternative. In addition to potential wavelength conversion and increasing the optical reach of the signal, OEO can provide other functions too. For example, in a network with subrate traffic, it is essential to bundle multiple connections together to better utilize the capacity of a wavelength. This bundling process is most effective when the traffic can be groomed at various nodes in the network. The grooming process is typically performed in the electrical domain using OEO conversion. The impact of physical impairment and features of the electrical layer on constrained routing is investigated in [46].

Most of PLI-RWA algorithms do not take into consideration the wavelength conversion capability. The few ones that allow for such a feature deal with a translucent network scenario and sparse regenerators that are capable of wavelength conversion [9,29,37,43]. The term sparse in this case means that the wavelength conversion is available only in selected nodes. In case that the node allows for sharing of wavelength converters between different input and output ports the number of conversions performed at the same time may be restricted.

The usual assumption is that full wavelength conversion, i.e., from any input to any output wavelength is available. On the contrary, wavelength converters with a limited conversion range allow an incoming wavelength to be switched only to a small subset of outgoing wavelengths [45]. To the best of our knowledge, the problem of the limited conversion has not been addressed extensively in the literature.

An interesting problem arises in the translucent network scenario and it concerns the optimization of placement of wavelength conversion-capable regenerators. Here the objective is to minimize the connection blocking probability resulting from both physical and network layer constraints. Such a problem was addressed in [63,65] by using some traffic-prediction-based heuristics.

### 3.4. Resilience and protection

In a transparent (and to some extents translucent) optical network, the impact of a failure propagates through the network and therefore failure cannot be easily localized and isolated. The huge amount of information transported in optical networks makes rapid fault localization and isolation a crucial requirement for providing guaranteed quality of service and bounded unavailability times. The identification and location of failures in transparent optical networks is complex due to three factors: (a) fault propagation, (b) lack of digital information and (c) large processing effort. Main challenges of fault localization in transparent optical networks include the selection of performance parameters to cover the full range of faults while ensuring cost effectiveness and preserving transparency. The placement of monitoring equipment to reduce the number of redundant alarms and to lower the capital expenses, and the design of fast localization algorithms are among challenges of fault localization in transparent optical networks.

During our literature survey we also noticed that very few works have addressed the issue of resilience and protection in the PLI-RWA algorithms. Authors in [54] propose an approach, which in addition to physical layer impairments and traffic condition, also takes into account the path reliability in the framework of a constraint based path selection algorithm. In [32] the effect of physical layer impairments on dedicated path protection schemes is investigated. The performance of dark and lit backup (protection) path is investigated and it is concluded that lit backup scenario introduces significant penalties in terms of blocking probability and vulnerability to failures. In the framework of all-optical networks with various path protection schemes, authors of [38] have proposed algorithms that exhibit low blocking probability without high computational complexity. The authors conclude that considering the blocking probability and required processing time, their dark backup algorithm performs better than lit backup; however lit backup path eliminates the need for signaling and enables faster network recovery. In [29] authors have proposed the Suurballe algorithm in the layered network graph, in order to find the shortest cycle passing through source and destination and using disjoint nodes. This cycle is then divided into primary and protection path. If no disjoint paths exist, the cycle having the minimum number of common vertices is selected. The work in [37] addresses the issue of survivability in optical mesh networks considering optical layer protection and realistic optical signal quality constraints. Three kinds of resource sharing scenarios, including wavelength-link sharing, regenerator (i.e. OEO) sharing between protection lightpaths, and regenerator sharing between working and protections paths are investigated in this work. In addition to an ILP-based solution, the authors have also proposed a local optimization heuristics approach and a tabu search heuristics to solve this problem.

### 3.5. Performance metrics

The traditional way of evaluating the performance of the proposed PLI-RWA algorithms in the literature has
been: the percentage of blocked connections versus the traffic load for dynamic traffic; and for static traffic, the same metric is also reported, as well as the amount of necessary resources (fibers, wavelengths, regenerator, etc.).

In a dynamic scenario, the network is designed and the objective is to route the maximum number of connections. For this reason, the percentage of blocked connections (or blocking probability) is used in order to compare the performance of different PLI-RWA algorithms.

On the other hand, for a static scenario we can find two possible cases. If the resources of the network are fixed, i.e. number of fibers, wavelengths, regenerators, etc., then the objective is the same as in a dynamic scenario: to route the maximum number of connections. But however, PLI-RWA algorithms with static traffic are usually utilized in the design phase of the network. In this case, the objective may be to minimize the number of necessary wavelengths or even, in the case of translucent networks, to minimize the number of required regenerators.

Few works report different metrics. For example in [35], authors evaluate their proposal presenting results of regenerator usage and number of necessary transponders. In addition to blocking probability, in [43] we find results of resource utilization (in terms of average and standard deviation of link, transmitter, receiver and electronic interface utilization). Moreover authors report results of computational time of the proposed algorithm. In [9], authors evaluate their proposed LERP algorithm using different metrics. They present usual results of percentage of blocked connection and required number of regenerators, as well as results of lightpath channels used per demand and regenerator repartition in the network. The work in [29] evaluates the performance of its proposal in terms of computational time, and dimensioning results (i.e. number of required fibers and regenerators). In [37], the computational time (in terms of running time of the algorithm), and also the number of required OEO modules and wavelengths are reported. Finally, in [59] the impact of crosstalk accumulation on the maximum transmission distance, blocking probability and BER (average value and distribution) and also fairness of the proposed algorithms in terms of blocking probability and BER are reported. A few other papers, consider the lightpath establishment setup time as a performance indicator. For example [70] addresses the extension of control plane considering the PLI constraints.

3.6. Evaluation of PLI-RWA algorithms

Another metric in this survey is the performance evaluation of the proposed PLI-RWA algorithm. Most of the papers have evaluated their proposed algorithms using simulation studies. However we have also found some experimental and analytical approaches that are proposed for performance evaluation. Table 6 presents the summary of various evaluation techniques that have been considered in surveyed works.

3.7. Centralized vs. distributed approach

Two different approaches can be followed to solve the PLI-RWA problem [21,66,67]. In the centralized approach, a single element stores the complete information of network topology, resource availability, and PLI performance in a central repository. This element is therefore in charge of collecting and updating all this information and also responsible for computing the optimal routes guaranteeing and satisfying the specific set of lightpath requirements such as optical signal quality, latency, etc. The central element could be either the network management system (NMS) or a path computation element (PCE) [66].

In the distributed approach, each node is responsible to compute, setup, and maintain lightpaths using a common and distributed control plane [21]. The nodes can collect the information on the status of the resource availability by means of a routing protocol, execute a RWA solver, and establish the lightpaths by means of a signaling protocol. To include the PLI constraints in the RWA problem, some extensions are necessary to the current signaling and/or routing protocol.

Table 7 classifies some of the selected papers according to their approach to solve the PLI-RWA problem. The details of how to implement these approaches are discussed in Section 4.

4. Impairment aware control plane extensions

This section introduces briefly the role of control plane in wavelength switched optical networks and then we focus on the works that has been done to address and include physical layer impairments in the control planes.

The introduction of a control plane (CP) is recognized as a necessary requirement for fast and flexible resource provisioning, easy network operation, and high reliability and scalability. The standardization process for such a control plane is currently being done independently by two different bodies: the automatically switched optical network (ASON) concept [68] developed by ITU and the generalized multi protocol label switching (GMPLS) suite of protocols [69] developed by IETF.

The main benefit of the ASON approach is the definition of the architecture, the requirements and the functionalities of the control plane independently of a particular choice of control protocol. Therefore, a variety of such protocols can be used ranging from the ATM family to MPLS and GMPLS ones.
Contrarily, GMPLS focuses on the implementation of the control plane, involving signaling (RSVP-TE), routing (OSPF-TE), and resource management (LMP) functions and protocols.

Although both GMPLS and ASON approaches for CP in optical networks are relatively mature and key standards are already available, they do not include any information related to physical impairments and thus are unaware of

Fig. 5. Physical layer impairment aware control plane extensions (OCC is the optical connection controller, and OXC is a typical optical cross connect). (a) PCE model; (b) routing model; (c) signaling model.
quality of optical signals. Some recent works deal with the problem of encompassing the PLI constraints into the GMPLS CP functionalities. Three different models have been proposed (see Fig. 5), namely the path computation element (PCE) model, the signaling model and the routing model.

4.1. PCE model

This model (Fig. 5a) has been first proposed in [70] and is based on the computation of the PLI-RWA problem in a centralized way utilizing a PCE element. PCE is defined in [71] as an entity that is capable of computing a network path or route based on a network graph, and of applying computational constraints during the computation. Its aim is hence to perform complex centralized route computation on behalf of the control module of the nodes.

In such a model, the PCE stores two databases. The first one is the traffic engineering databases (TEDs), which are located in the nodes. The second one is the centralized TED, which is updated from former TEDs through a standard, distributed routing protocol. The other database is the PLI database (PLID) obtained by the network management system (NMS) or through a performance monitoring system. The PLID maintains up-to-date information on any possible PLI concerning any network link. Whenever a new connection request arrives to a node, it sends a query to the PCE. The PCE computes the required path taking into account the TED and PLID information and sends back the computed explicit route to the source control module. The source node, using the standard signaling protocol (PATH/RESV messages), establishes the lightpath.

Slightly different approaches are followed in [72,73]. In [72] PCE interworks only with NMS. In such a case, the NMS is in charge of collecting the connection requests, send them to the PCE together with all required information such as PLI performance, topology and logical link status, and get the computed routes. Afterwards, NMS sends the routes to the nodes which establish the lightpaths using the standard signaling protocol.

In [73], the authors propose an OPM manager, which is directly integrated in the NMS instead of a separate PCE, but the behavior and functionalities are similar to the PCE model.

4.2. Routing model

The routing model (Fig. 5b) has been mentioned in [21] and consists of extending the routing protocol (as e.g., the OSPF-TE in GMPLS CP) to involve the PLI constraints into the PLI-RWA problem. As described in [66], each node is in charge of storing updated TED and PLID databases on the resource utilization and on the PLI performance concerning any link in the network. As in the case of the TE attributes, local PLI (i.e., PLI performance of local node and of the attached links) can be included in the PLID using local monitoring while remote PLI can be obtained by exploiting the extended routing protocol. Whenever a new connection request arrives to a node, an on-line PLI-RWA algorithm computes the route taking into account the TED and PLID information. Once the path is computed, the node activates the standard signaling protocol to establish the lightpath.

A different routing model is considered in [56], where no routing protocol extensions are required. Global wavelength availability is stored in the TED of the nodes and updated by the standard routing protocol while the PLID is not necessary. Indeed, only PLI constraints of static nature or function of the number of active wavelengths are considered and their mathematical models are preloaded into the control module of the nodes. Any incoming connection request triggers therefore the PLI-RWA algorithm that computes a set of candidate routes according to the TED information and checks its feasibility by means of the mathematical models. If at least one feasible computed route exists, a lightpath is then established by the signaling protocol.

4.3. Signaling model

The signaling model (Fig. 5c) has been proposed in [52] and consists in extending the signaling protocol (as e.g., RSVP-TE in GMPLS CP) to encompass the PLI constraints. In such a model no modifications are introduced in the routing protocol [56,74]. Whenever a connection request arrives to a node, it computes a route according to the TED information and launches a setup request message in the network. This message collects estimated PLI performance of any traversed link between source and destination node. In fact, each node must store updated local PLID and mathematical models to calculate the PLI performance. If the accumulated PLI performance on the receiver interface at the destination node is compliant with an acceptable signal quality, a positive response message is sent back to the source node and the lightpath is established. If the accumulated PLI performance is not satisfactory, an error message is sent back to the source node and another attempt can be triggered following different routes.

In order to decrease the delay of the lightpath establishment process, some improvements are described in [30,75]. In [30], four different approaches are proposed and compared: in K-seq, the source node computes k different routes and sequentially attempts to establish a feasible lightpath; in K-par, the source node sends simultaneously k setup messages and the destination node can pick one according to some criteria and generates the response message; in HbH, the route is computed hop-by-hop, which means that each node takes into account only the information on the adjacent links; in FF, the setup message is flooded to the entire network. In [75] the lightpath provisioning with signaling feedback (LPSF) concept is defined. It exploits the error message delivered to the source node adding some feedback information on the PLI performance of the rejected route. The source node can therefore store this information in the local PLID and compute additional feasible routes.

4.4. Comparison and discussion

Being a centralized approach, the PCE model is able to provide optimal path computation in terms of both network utilization and optical signal quality. It also does not
require any modification or extension to the current signaling and routing protocols. At the same time, PCE has a global view of the network and, if there is no inconsistency in the databases, the setup procedure does not require any re-attempt, which can speed up the service provisioning. Nonetheless, it suffers from scalability problems and in case of failure rapid restoration cannot be achieved. There are some reasons to consider the PCE model for multi-domain scenario, where an abstraction of the entire routing area may optimize the inter-domain paths [66,70].

Maintaining accurate routing information on all network nodes under dynamic traffic is extremely difficult. Therefore, routing model seems less advantageous solution since in addition to the TED information it requires the global dissemination of the PLI performance data. It may give some benefits in case of static traffic, or less volatile traffic conditions, but in such a case the PCE model may outperforms the routing one.

The signaling model seems to be the easiest and fastest way to encompass the PLI performance into the RWA problem. On the other hand, it is not able to provide optimal resource utilization and signal quality. It may require high setup delay due to the re-attempts of failed lightpath establishment processes (see Table 8).

5. Discussion

In the previous sections of this paper we have reported and overviewed some of the most relevant work in the literature related to the PLI-RWA algorithms and the required modification to the control planes. However there are still some points which are either not reported in the literature or few works are devoted to them.

Multicast routing in wavelength switched optical networks has received some interest recently. Data duplication is performed in the optical domain at a set of branching nodes by splitting the optical signal using passive splitters. Considering multicasting in RWA problem is also NP-complete like classical RWA [76]. Very few works addressed the multicast problem taking into account the physical layer impairments and, in particular, they only consider the power loss due to the passive splitters as a constraint [51,77–79].

In general, the QoS support in physical impairment-aware optical networks has two folds. The first one corresponds to the physical layer performance and it refers to the pre-defined level of signal quality, as measured at the destination receiver, that allows for flawless network operation [39,60,36]. Another interpretation of QoS can be found in [54]. There the authors introduce a model of network that is able to support differentiation of lightpath requests according to a set of routing constraints (such as e.g., max. transmission quality degradation, max. delay, or reliability).

Although the PLI-RWA algorithms presented in the literature focus mainly on the optical circuit-switching (or wavelength-routed) networks, still there are issues specific to the optical burst/packet switching networks (OBS/OPS) that have to be addressed. To support short connection holding times it was proposed to incorporate the information about physical impairments into a setup control message, by means of a data vector that is processed at consecutive nodes [80]. Another problem concerns OBS networks, where the transmission offset time may be affected by optical signal dispersion effects [81].

Few works in the recent literature address the problem of regenerator and monitoring equipment (optical performance or impairment monitoring) placement and allocation. The regenerator placement is addressed in [9,13,24,29]. The allocation of regenerator is proposed in [35,43,54]. In [43], the efficient regeneration-aware algorithm minimizes the number of used regenerators along the selected lightpath as well as the PLI constraints. In the DWP (distributed discovery of wavelengths Paths) method [54] one of the objectives is the minimalization of the utilization of electronic regeneration.
6. Conclusions

A comprehensive survey that covers the incorporation of physical layer impairments in planning and operation of translucent/transparent optical network (i.e. PLI-RWA) was presented in this work. Physical layer impairments can be classified into linear (i.e., attenuation, CD, PMD, FX, crosstalk, ASE noise, insertion loss, and PDL) and non-linear (i.e., SPM, XPM, FWM, SBS, SRS) effects. Analytical models (e.g. Q-Factor) or a hybrid approach considering analytical, simulation and experiments are proposed for modeling the physical impairments and incorporating their impacts in RWA algorithms.

Heuristics, meta-heuristics and optimization techniques are proposed as algorithmic approaches to solve PLI-RWA problems. The general approach to address the PLI-RWA problem can be divided in two main categories. The first trend utilizes traditional RWA algorithms and after selecting the lightpath the physical constraints are verified; in this approach the PLI-RWA algorithm is not deliberately designed for routing with PLI constraints. The second approach is to use some metrics, which are related to the PLI constraints as cost of the links in order to compute the shortest path(s). Assuming good algorithms are in place, a small number of wavelength conversion (either via OEO or all-optical conversion) is needed to approximate the performance of opaque network architectures. In addition to wavelength conversion via OEO conversion, grooming will become available in PLI-RWA algorithms and network planning decisions in general.

The physical impairments and PLI-RWA algorithms can be incorporated in control planes using PCE, routing and signaling model. The PCE model is able to provide optimal path computation in terms of both network utilization and optical signal quality. It also does not require any modification or extension to the current signaling and routing protocols. PCE has a global view of the network, which can speed up the service provisioning. Nonetheless, it suffers from scalability problems. Routing model seems less advantageous solution since in addition to the TED information it requires the global dissemination of the PLI performance data. The signaling model seems to be the easiest and fastest way to encompass the PLI performance into the RWA problem. On the other hand, it is not able to provide optimal resource utilization and signal quality. It may require high setup delay due to the re-attempts of failed lightpath establishment processes and possible sub-optimal route decisions due to impairment-unaware route computation.

There are quite few proposals that address the resilience and protection issues in the PLI-RWA algorithms. In addition to simulation studies, experimental and some analytical models are available to evaluate the performance of the proposed algorithms. None of the surveyed works considers the inaccuracy of the physical impairment information (analytically computed or measured) into their PLI-RWA algorithms. The proposed adaptive PLI-RWA algorithms simply change their decisions assuming that the physical information are completely accurate.

Regenerator and/or monitoring equipment placement are important factors in the design phase of the network. By using a proper regenerator or monitoring equipment placement strategy in some nodes of the network, it is possible to obtain similar performance (in terms of blocking probability) of an opaque network with much lower cost. However, this topic is not enough investigated in the literature.

The overall conclusion is that PLI-RWA algorithms play important roles in maximizing the performance of an optical network design. These algorithms, when exploited in transparent or translucent networks planning and operation tools, can provide similar utilization as an opaque architecture, but with lower cost.

Acknowledgements

The work described in this paper was carried out with the support of the BONE-project ("Building the Future Optical Network in Europe"); a Network of Excellence funded by the European Commission through the 7th ICT-Framework Program and the National Spanish project CATARO (TEC2005-08051-C03-01). The authors would like to thank the EC FP7-DICONET project for funding this work. The authors would also like to thank Dr. Yvan Pointurier for his comments and suggestions.

References


Prof. Josep Solé-Pareta obtained his M.Sc. degree in Telecom Engineering in 1984, and his Ph.D. in Computer Science in 1991, both from the UPC. In 1984 he joined the Computer Architecture Department of UPC. Currently he is Full Professor with this department. He did a Postdoc stage (summers of 1993 and 1994) at the Georgia Institute of Technology. He is co-founder of the UPC-CCABA. His publications include several book chapters and more than 100 papers in relevant research journals (>20), and refereed international conferences.

His current research interests are in Autonomic Communications, Traffic Monitoring and Analysis and High Speed and Optical Networking, with emphasis on traffic engineering, traffic characterization, MAC protocols and QoS provisioning. He has participated in many European projects dealing with Computer Networking topics.

Dr. Ioannis Tomkos, has the rank of Full Professor at Athens Information Technology Center, serves as its Associate Dean (since 2004) and is an Adjunct Faculty at the Information Networking Institute of Carnegie-Mellon University, USA. In the past (1999–2002) he held a senior scientist position at Corning Inc. USA. He joined AIT in 2002 where he founded and serves as the Head of the “High Speed Networks and Optical Communication (NOC)” Research Group that participates in many EU funded research projects (including five running FP7 projects) in which Dr. Tomkos is representing AIT as Principal Investigator and has a consortium-wide leading role (e.g. Project Leader of the EU ICT STREP project DICONET, Technical Manager of the EU IST STREP project TRIUMPH, Chairman of the EU COST 291 project, WP leader). Dr. Tomkos has received the prestigious title of “Distinguished Lecturer” of IEEE Communications Society for the topic of transparent optical networking. Together with his colleagues and students he has authored more than 200 peer-reviewed articles and his work has received over 600 citations. Dr. Tomkos has served as the Chair of the International Optical Networking Technical Committee of IEEE Communications Society and a member of the IEEE ComSoc’s Technical Activities Council. He is currently the Chairman of the OSA Technical Group on Optical Communications and the Chairman of the IHP working group on “Photonic Networking”. He has been General Chair, Technical Program Chair and member of the steering/organizing committees for the major conferences (e.g. OFC, ECOC, IEEE GlobeCom, IEEE ICC, etc.) in the area of telecommunications/networking (more than 50 conferences/workshops). In addition he is a member of the Editorial Boards of the IEEE/OSA Journal of Lightwave Technology, the OSA Journal of Optical Networking, the IET Journal on Optoelectronics, and the International Journal on Telecommunications Management.