Aalborg Universitet

AALBORG UNIVERSITY

DENMARK

A Methodology for Physical Interconnection Decisions of Next Generation Transport Networks

Gutierrez Lopez, Jose Manuel; Riaz, M. Tahir; Madsen, Ole Brun; David, Mathieu; Pedersen, Jens Myrup

A Methodology for Physical Interconnection Decisions of Next Generation Transport Networks

Jose M. Gutierrez, M. Tahir Riaz, Ole B. Madsen, Mathieu David, Jens M. Pedersen

Department of Electronic Systems, Networking and Security Section, Aalborg University, Denmark
Fredirk Bajers Vej 7, 9220 Aalborg , Denmark

Emails: jgl, tahir, obm, md, jens@es.aau.dk

The physical interconnection for optical transport networks has critical relevance in the overall network performance and deployment costs. As telecommunication services and technologies evolve, the provisioning of higher capacity and reliability levels is becoming essential for the proper development of Next Generation Networks. Currently, there is a lack of specific procedures that describe the basic guidelines to design such networks better than "best possible performance for the lowest investment". Therefore, the research from different points of view will allow a broader space of possibilities when designing the physical network interconnection. This paper develops and presents a methodology in order to deal with aspects related to the interconnection problem of optical transport networks. This methodology is presented as independent *puzzle* pieces, covering diverse topics going from novel design criteria to well-known organized topologies. These can be used to investigate the influence of the physical interconnection of networks over their performance properties, and draw conclusions to improve the current decision support techniques related to this theme. In addition, several examples of the use of this methodology are presented.

Keywords: Network Planning, Physical Interconnection, Optical Infrastructure, Transport Networks, Heuristics.

1. Introduction

The fast evolution of communication technology and services is causing a gradual increment of the bandwidth and reliability requirements to be fulfilled by the network infrastructure [1]. At the same time, our daily lives are becoming highly dependent, in both social and professional contexts, on the technological features of our telecommunication devices.

In order to be able to support all the traffic and quality of service demands, there is a need of high performance transport systems, and focusing on the specific field of this work, **a high performance optical backbone infrastructure** [2]. Failure to achieve these challenges may lead to significant profit losses or the physical infrastructure to become the bottleneck for future telecommunication developments [3]. In this paper, network infrastructure is referred to as the physical interconnection of nodes by optical links placed in trenches along roads.

The deployment of optical networks is a long and expensive process, due to the trenching tasks involved, especially for large geographical areas such as national or continental territories. Moreover, the lifetime of the physical infrastructure is rather

2

long, up to 40 years [4]. These features make such networks deployment a long term investment project, which should be carefully planned. This high investment is required for the proper development of the performance of Next Generation Networks (NGN). In addition, when combined with a **reliable**, **preventive** and **green** planning, a better outcome can be achieved [5].

Optical network systems are very complex; each of the network layers can have great impact on the overall performance. In relation to backbone networks, the infrastructure can be limiting the whole performance of a network just by the fact that the physical interconnection scheme has not been carefully planned [3]. In fact, when the network requires some kind of physical upgrade, the solution's costs in economic terms might be much more significant, due to trenching and deployment tasks, than at higher layers where software updates might be enough to solve the problem [6].

Therefore, planning is crucial and even small improvements have high economic impact nowadays. However, no well documented methods exist for the whole interconnection planning process leaving room for the development of this work.

The main overall challenge can be described as the physical interconnection decision problem of a set of nodes. This interconnection can be configured in many different ways, and several of these might be "optimal", depending on the objective function and constraints of the optimization. For example, for the interconnection problem of a number of nodes forming a national backbone, some of the most common network planning tasks and decisions are:

- Decide the interconnection scheme. Which nodes are directly physically connected.
- Define failure support mechanism such as dedicated or shared protection.
- Dimension the links, how much fiber is required to support the given traffic demands including the spare capacity required for failures situations.
- Define the main constraints or objectives of the deployment, such as budget limitations or graph's minimum cut; among others.

The combinatorial complexity nature of the problem, combined with the decision support based on the balance of certain parameters (i.e. cost vs. availability or cost vs. energy efficiency), makes the research on the physical interconnection in optical backbone networks an interesting non-trivial focus for engineers. For more information about the complexity in graphs is recommended to read [7] and [8].

The **main goal** of this work is to present a network planning research methodology in order to be able to establish the rules and procedures for future optical transport networks interconnection design. The network infrastructure topic has not yet been covered in depth in a scientific research context. There are many contributions, [9], [10], [11], or [12] among others, that marginally cover or solve specific problems, but to the best of our knowledge, the relevant literature does not include an overall well-defined methodology. This might be beneficial for research purposes,

in order to draw conclusions regarding the influence and relevance of the physical network interconnection in transport optical networks. The research framework in this work is referred to as the analytical developments, experiments, empirical results, and conclusions that contribute to define rules and procedures for interconnection decision.

The rest of the document is organized as follows: Section 2 covers the current situation of this type of networks, definitions of important concepts, and summary of the notation used in this document. Section 3 summarizes the overall proposed methodology. Section 4 introduces the pre-calculated information required as an input to the problem. Section 5 presents the different parameters associated with the solutions to the problem. Section 6 covers the complexity analysis, search methods, and objective functions. Section 7 presents experiments to illustrate the use and application of the methodology. Section 8 presents the conclusion of this work.

2. Background

This section presents the background and additional information necessary for a more complete understanding of the work. The current situation of transport networks infrastructure and some relevant definitions are described in the following paragraphs.

2.1. *The Present*

Transport networks infrastructure is usually formed by a number of nodes connected as a mesh or ring mesh. When these are deployed, fiber links might be disjoint from each other, but they are deployed sharing the same trenches and ducts. Hence, when a failure occurs, more than one fiber link can be affected, causing major services disruption. For example, in the beginning of 2008, the cable connecting Europe and Middle East suffered four cuts, affecting millions of users [13].

In relation to the distribution of the links, currently, it is widely accepted that these ring interconnections are reliable enough for the demands of the users. However, if the current evolution of telecommunication demands keeps heading towards a more IT dependent society, higher degree physical networks (> 2) can significantly contribute to the improvement of failure support, congestion control or delay propagation aspects.

Therefore, new perspectives are required in order to design Next Generation Backbone Networks. As it is discussed in depth in Section 6, the complexity of the physical interconnection problem makes it infeasible to follow the exhaustive search design approach. A scientific effort is required in order to improve the current decisions support procedures in terms of solution quality and processing time. The physical interconnection decision problem can be simplified to acceptable complexity levels by concisely defining new research goals and evaluation procedures, as it is presented in the following sections.

4

2.2. *Preliminary Concepts*

The following key terms are used along the document:

Physical interconnection: Refers to the way the nodes are interconnected by fiber, including trenches. The nodes are associated to a geographical location and the links are characterized by their physical length. Links are considered unidirectional and, in case of assuming symmetric traffic, fibers are equally distributed in both flow directions.

Topology: Refers to any specific graph the network links follow. It can be described using the adjacency matrix.

Links disjointness: Implies that all physical links are disjoint from each other. Two different links cannot share the same trenches.

Hop: Defines the trip of a data packet from one node to its immediate neighbor.

Length and Distance: Define two different concepts in this paper. Length is always used in physical terms (km) and distance is always used in transmission path terms (hops). For example, the path length between A and B is 304 km, and the path distance between A and B is 5 hops.

Minimum Cut: Refers to the minimum partition of the nodes in a graph which removal results in two disjoint subsets.

Connectivity factor: Defines the density of the number of links L in relation to the number of nodes N, defined by $\alpha = \frac{2 \cdot L}{N \cdot (N-1)}$.

k redundant paths: Refers to the existence of k disjoint paths between a source and destination pair.

2.3. *Notation Summary*

Table 1 lists in alphabetic order the notation used along the document.

3. Summary of the Proposed Research Methodology

This section introduces and summarizes the concepts behind the proposed methodology in this work. Methodology refers to the set of techniques used in order to investigate the role of the physical interconnection in optical transport networks.

This methodology must be defined in order to provide guidelines about how to deal with different scientific problems regarding the physical interconnection of networks. Some of the concepts have been marginally applied to network-related problems that are commented along the document, but a global coverage of this interconnection decision has not been proposed in the relevant literature to the best of our knowledge.

The methodology can be divided into modules, each covering a field in the context of this problem. Hence, this methodology can be characterized by its modularity. These modules are independent from each other in the sense that they can be modified or extended without interfering with the others. This allows simplifying the process of expanding or updating the methodology along with future changes in the

Table 1.   Notations

α	Connectivity factor	λ	Wavelength
A_x	Availability variables	C^x	Adjacency matrix
CW	Capacity per wavelength	d	Destination node
E_x	Emissions variables	EC_x	Economic cost variables
G	Graph	Hp_i	Path's Distance (Hops)
I_x	Costs variables	k	number of disjoint paths
L	Numbers of links	l_{nt}	Networks length
Ld	Link traffic load	Lm	Length matrix
Lp_i	Path's Length (km)	MAX	Maximization
$Max(Sol_x)$	Max. number of solutions	MFT	Mean Failure Time
MIN	Minimization	$min^k(f(x))$	k-lowest result for $f(x)$
$MTTF$	Mean Time To Fail	N	Number of nodes
nf	Number of fibers	P_h	Set of paths with max. dist. h
p_i	Path	Pop_x	Population variables
s	Source node	S_L	Set of links
S_N	Set of nodes	TRF	Traffic
$UserBW$	Traffic volume per user	W	Wavelengths per fiber
Wgh	Weight		

network requirements, avoiding important or complex modifications.

This methodology is presented in *chronological order* in relation to the problem solving process. Three main types of modules are covered: Pre-calculated Information in Section 4, Characteristic Parameters in Section 5, and Search and Evaluation in Section 6.

The research topics that can be dealt with by this methodology are very diverse. The main scope is to research on different points of view to design and deploy networks better than the "*lowest possible budget approach*" [2]. One of the main ideas is to identify how the physical interconnection itself can affect the availability of connections, traffic distribution or deployment expenses. Interconnection schemes that follow regular or "organized" graphs, have been widely studied regarding logical or virtual topologies from a graph theory perspective, for example in [14]. Similar structures can also be used in order to physically interconnect nodes taking advantage of their graph properties. For example, a degree 3 topology can support up to two simultaneous failures without loss of connectivity, regardless the failure element.

Besides the overall global objectives covered by this methodology, each module and each of their options have specific purposes in specific experiments. The term experiment is used along the document, and it generically refers to any empirical test or exercise performed in order to obtain conclusions regarding a specific hypothesis.

For a single problem, not all of the modules are required, and in order to properly identify how to deal with it, each module has to be individually defined. Each subsection in Sections 4, 5 and 6 corresponds to a module implicitly covering the following aspects: Introduction, related work, purpose, and options when relevant.

The modules and their groups are schematically presented in Fig. 1. This figure is represented following the chronological order of the problem solving. Firstly, the

6

pre-calculation information is required. This information may be part of the input to the problem or should be determined. In the second case, it can be easily processed, only once, to provide or build the proper scenario. Secondly, this information is forwarded to an iterative process (or loop) where the steps *"parameter calculation-evaluation-search"* are repeated until a satisfying solution is achieved.
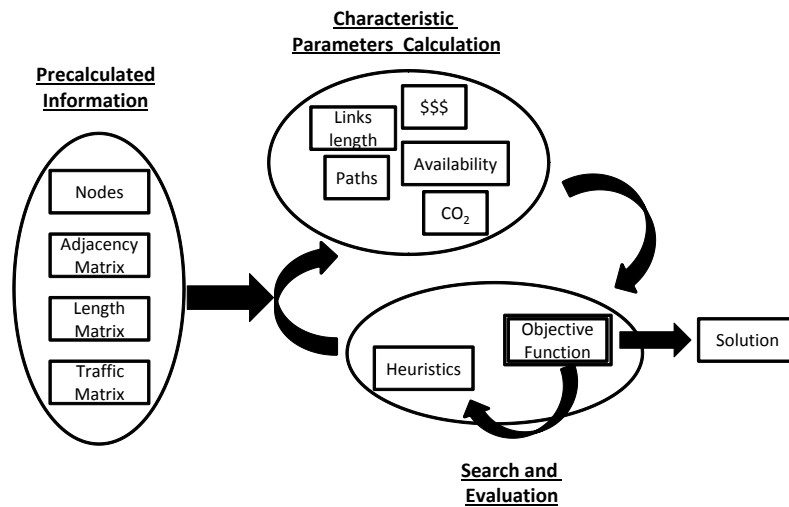

Fig. 1.   Modules Diagram

4. Pre-calculated Information

This section covers the modules of the methodology related to the input information to the problem solving process. The relevance of each of the parameters is dependent on the specific problem to be solved and they should be defined according to its nature. These modules are: Nodes Location and Length, Adjacency, and Traffic Matrices.

4.1. *Nodes Location*

The location of the nodes to interconnect contributes to the definition and conclusion of the different kinds of experiment results. It is very important to choose the proper relative location of the nodes from each other in order to study the properties of the different interconnection schemes. The location may follow different approaches depending on the kind of performed experiments and their objectives. These objectives are very diverse such as efficiency test of search methods, proper-

ties generalization of different deployed topologies or simply problem solving using more commonly-known scenarios with real cities of well-known research networks.

The nodes location options defined in this methodology are the following:

Schematic:

The nodes are located in such a way that the resulting interconnection is shaped as a pre-defined known solution, see examples in Fig. 2. This idea can be used to test the search algorithms efficiency, or determine the convenient numerical parameters for a specific network size. Since the solution of the optimization is known, i.e. shortest physical network length, the accuracy or error rate of search algorithms can be determined to be later extrapolated to real interconnection problems. For example, in the case of dealing with a ring topology, if the nodes are located following a circle line, the shortest solution is exactly that circle. In a similar way this idea can be extrapolated to other organized topologies.
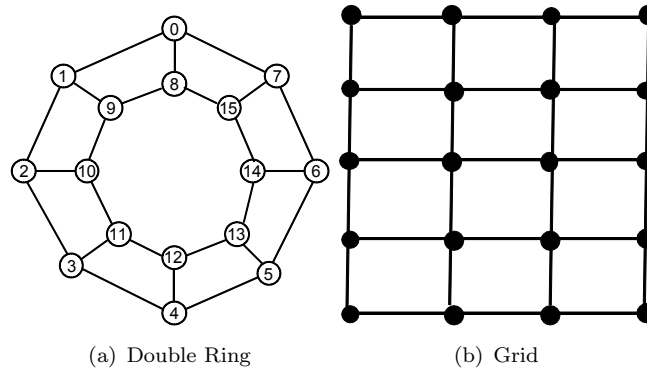

(a) Double Ring   (b) Grid

Fig. 2.   Schematic Examples

Random:

In this case, the specific location of the nodes is not relevant. It can be used for experiments where empirical data is required, covering hundreds or even thousands of samples. For example in [15], the hypothesis was that the length of a network following an organized topology (length optimized) is somehow proportional to the area covered by the nodes. To test it, hundreds of sets of randomly placed nodes networks were length optimized to collect enough samples. Results show that the length of the network can be predicted within a certain error that is variable depending on the topology. These predictions were applied to a real geographical area to validate the conclusion.

This example is summarized in Section 7 as part of the presentation of the practical use of this methodology.

Traditional Research Networks and Real Geographical Information:

Many studies deal with optimization problems or network analysis on fixed physical infrastructure. For example in [11] or [12], experiments are carried out over real network graphs such as *Arpanet*, *UKnet*, *EON* or *NSFNET* to minimize the number

8

of wavelengths or the capacity required for given traffic demands.

The use of these networks for experiments is convenient in order to keep results and conclusions in known and widely accepted scenarios by the scientific community.

Also, scenarios using real cities locations and relief conditions provide a more reader-friendly environment for presenting experiments and conclusions.

4.2. *Length Matrix*

One of the most important parameters is the physical length of the potential links to be used for the interconnection of nodes. This information can be precomputed to be later applied to each specific potential solution in order to calculate the relevant characteristic parameters described in Section 5.

The Length Matrix contains information regarding all the potential links in the network (all-to-all) and is formally defined by:

Definition 4.1. Let S_N be the set of nodes to be interconnected, $N = Size(S_N)$ being the number of nodes. Let Lm be a NxN matrix containing the information about the potential links length from all-to-all nodes. Each element Lm_{ij} corresponds to the length of the link between nodes i and j belonging to S_N.

There are different ways of approximating the value of each element Lm_{ij}:
Weighted Euclidean length:
Euclidean length is the most commonly used approach when just some distance values are required to run experiments, but finding these is not part of the problem.

In addition, a weight factor Wgh can be used as the ratio *Euclidean/Real* length in order to get a more realistic approach, since fiber is usually deployed along the roads and not on a direct straight line. As a rule of thumb, a commonly accepted approximation factor is $Wgh = \sqrt{2}$, but more exhaustive studies such as [16] have done more accurate approximations using statistical and mathematical methods. The formal definition of the link's length between nodes $i - j$ characterized by Euclidean coordinates (X, Y) is presented in Eq. (4.1).

$$Lm_{i,j} = Wgh \cdot \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2} \qquad (4.1)$$

A variation of this method, especially useful for large areas such as the US territory, is to calculate the physical length between points on a map using the *great circle* concept that considers the latitude and longitude of the end points and the curvature of the Earth. The Haversine formula is used to find its value along the surface of a sphere [17].
GIS Tools:
Geographical Information Systems and road databases are very helpful when dealing with real scenarios, since roads are commonly used for the fiber deployment. Currently, there are many online GIS based applications that provide information about the roads to follow between two points on a map. The extraction of this information is rather fast and simple; a short script can be implemented in order

to automatically obtain this information. This method does not guarantee that the given roads will be the ones used for the implementation, and it does not ensure that the different links will be physically disjoint from each other, but it is a good approximation to work with.

4.3. *Adjacency Matrix*

The adjacency matrix represents the way nodes are interconnected to each other. The structure of an adjacency matrix C^{Top} for a topology Top with N nodes is described in Eq. (4.2).

$$C_{ij}^{Top} = \begin{cases} 1 \text{ if } i \Rightarrow j \\ 0 \text{ rest} \end{cases} \quad 0 < i, j \leq N \tag{4.2}$$

If the links are unidirectional, the adjacency matrix must be symmetric. If there is a direct fiber link from i to j there is a direct fiber link from j to i following the same trench.

The adjacency matrix is part of the input information to the problems described in this work. In this paper, this type of problems is classified as *topology related approach*. However, this matrix can also be the solution of the problem, in this paper classified as *link related approach*. This concept and both approaches are discussed in depth in Section 6.2, but as a brief introduction to understand the idea, Fig. 3 illustrates a very simple generic example.
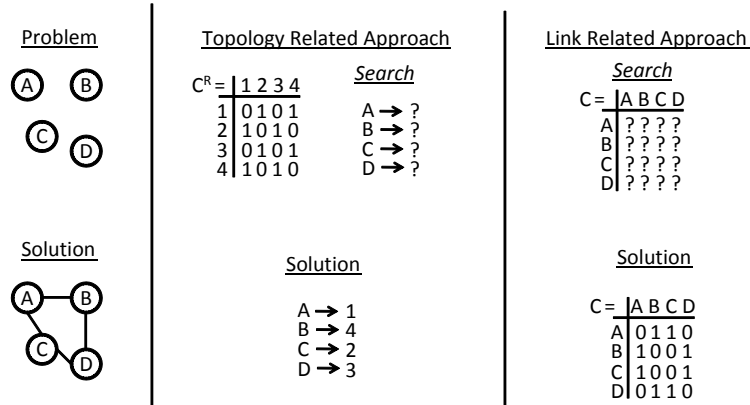
Fig. 3.   Problem Approach Example

Basically, the problem consists of interconnecting 4 nodes (A,B,C,D) as a 2-connected graph. In the *topology related approach*, the resulting interconnection must follow a concrete generic adjacency matrix, C^R, corresponding to a Ring. Hence this is a mapping problem between (A,B,C,D) and (1,2,3,4).

Instead, in the *link related approach*, the objective is the adjacency matrix C. It may represent any topology fulfilling the given constraints.

10

Based on the concrete solution given in Fig. 3, it is simple to identify the difference between both approaches with the result to the problem.

Adjacency matrices are used to mathematically represent and work with network topologies. In this work, these topologies are considered as organized due to their structural properties. These patterns can be easily identified in the adjacency matrix. Examples of these topologies are Double Ring, Chordal Rings, Honeycomb or Grid [18].

Network topologies have a significant influence on the traffic transactions between nodes. An increment on the connectivity factor in networks usually allows shorter path distances between pairs of nodes. This implies a reduction of the routed traffic at each of the nodes, and consequently the nodes' switching and links capacities are better utilized.

Following this assumption, the benefits can be depicted in two ways:

- Given fixed incoming access network traffic, less fiber and smaller switches are required for the backbone network to support the traffic.
- Given fixed switch sizes and fibers per link in the backbone network, higher incoming access network traffic can be supported.

4.4. *Traffic Matrix*

Traffic Matrices represent the information transactions between nodes in a network. Problems dealing with or requiring link dimensioning need information related to these transactions. Sometimes, this information can be extracted from real life networks, but usually these data is not revealed by operators.

In that case, the easiest solution is to randomly generate a traffic matrix, but there are other more realistic approaches. The kinds of problems requiring a traffic matrix are very diverse: Fiber cost optimization, path failure analysis or traffic patterns influence over the interconnection decision.

The next few paragraphs describe some of the most popular methods of creating a traffic matrix to work with.

Population Based Distribution:

Traffic matrices can be created based on the population covered by each node. It is assumed that the aggregated traffic at each node is proportional to its population. This traffic is distributed to the rest of the nodes proportionally to their population. This concept is formally described as follows.

Let $TRF(s,d)$ be the data transactions between a source s and a destination d nodes; $UserBW$ being the average bandwidth consumed per user (or inhabitant). $TRF(s,d)$ is defined by Eq. (4.3). This approach has been adopted, for example, by studies on Cost 239 research network [19]. The variables Pop_s, Pop_d and Pop_t are the population coverage by s, d, and the complete network respectively.

$$TRF(s,d) = UserBW \cdot Pop_s \cdot \frac{Pop_d}{Pop_t} \qquad (4.3)$$

The networks will be required to support future traffic under future conditions. Traffic is naturally increasing in time and the number of users may vary. Hence, all the variables (Pop and $UserBW$) values should be estimated for the end of a defined working period and then applied to calculate the matrix.

This traffic distribution leads to a symmetric matrix, and implicitly, each node's local traffic never reaches the backbone network, $TRF(s, s)$. This is the fraction of the assigned traffic to s that is not distributed to the rest of the nodes.

In case of distributing 100 % of the assigned traffic to each node to the rest of the network, the resulting matrix is not symmetric, i.e. $TRF(d, s) \neq TRF(s, d)$. Eq. (4.4) formally describes the concept.

$$TRF(s, d) = UserBW \cdot Pop_s \cdot \frac{Pop_d}{Pop_t - Pop_s} \tag{4.4}$$

Extensions of this approach can also consider external traffic flowing in and out the covered territory, i.e. international traffic. In this case, the concept of *node to gateway* traffic must be added to the regular *node to node* transactions.

In a near future, communication machine-to-machine will also significantly contribute to the network's traffic load. The consideration of not only human population but also machine population coverage is a possibility.

5. Characteristic Parameters

This section describes the modules corresponding to the calculation of parameters related to the solutions. Their numerical values are forwarded to the objective function to be evaluated. The search and problem solving is considered as an iterative process. In each of the iterations, a potential solution (or set of solutions) is provided, and its characteristic parameters must be individually calculated in that process. For example, the deployment cost, availability, or path distances can be considered as characteristic parameters.

5.1. *Network Length*

The network length l_{nt} can be calculated by the sum of the resulting elements of the product between the Length and Adjacency Matrices. The length of a network following a topology Top is formally defined by Eq. (5.1).

$$l_{nt} = \sum_{\forall \, i,j \, \in \, S_N} C_{ij}^{Top} \cdot Lm_{ij} \tag{5.1}$$

5.2. *Paths and Routing*

A specific routing scheme must be defined in order to be able to dimension any network. There are many possibilities and variances in order to deal with specific congestion, delay, reliability, or energy efficiency problems. In this work, static traffic matrices and routing are assumed for the required purposes.

The main problem when dealing with paths and routing is that they actively contribute to the optimization procedure, (objective function) and consequently, paths must be determined for each one of the evaluated solution in this process. Thus, in each iteration and solution, $N \cdot (N-1)/2$ primary paths must be calculated (assuming that the path from s to d is the same as from d to s). This number increases when alternative protection paths are also involved in the process. These operations often become a significant time consuming factor when performed for thousands of iterations.

In case of using a specific adjacency matrix, the complexity can be reduced by generically pre-computing a set of potential paths. Furthermore, if the topology is node-symmetric, the path calculations are only required from a single source node to later be extrapolated to the rest of the source nodes. Then, for each solution, the convenient k disjoint paths can be selected, k being the number of disjoint paths.

For example, based on the same simple idea shown in Fig. 3, these paths are generically precomputed with nodes (1,2,3,4). For each solution, (A,B,C,D) is assigned to (1,2,3,4) and the correspondent specific paths are determined by substitution. Finally, based on the requirements, the k disjoint paths are selected (only $k = 2$ in this case) to be used for other parameter calculations. Similar concepts are used for wavelength assignment optimization as described in [20].

Formally, the set of precomputed paths, $P_h(s, d)$, can be defined as the set of all the possible paths from a source node s to a destination node d, with a maximum hop distance of h.

In very specific cases, the selection of the first path(s) incurs in the impossibility of determining k disjoint paths. For example, if $k = 3$, the selection of the first two paths might not allow to find a feasible third one in $P_h(s, d)$. This problem must be avoided when the paths are selected.

The following paragraphs formally define the selection of the k disjoint paths for some of the most common static routing schemes, all based on $P_h(s, d)$.

Let $Hp_i(s, d)$ be the distance and $Lp_i(s, d)$ the length of each path $p_i(s, d) \in P_h(s, d)$. And let $P'_h(s, d) \in P_h(s, d)$ be a set of k disjoint paths.

Regarding notations, $min^k(f(x))$ refers to the k elements returning the lowest result for $f(x)$ and MIN refers to minimization.

Minimum Hop Routing, MH:

This type of routing is interesting to provide paths with the least number of nodes involved. This can be beneficial in e.g packet switching networks, where delays are mainly caused in the nodes by routing tasks. Less intermediate-nodes in a path might imply lower transmission delay. Also, the unavailability due to nodes failure can be minimized. The selection of k disjoint paths following this routing type is defined as follows:

Definition 5.1. $p_{i'}(s, d)$ is a selected disjoint path in $P'_h(s, d)$ if $Hp_{i'}(s, d) \in min^k(Hp_i(s, d))$ and does not share any element with any of the other paths in $P'_h(s, d)$, for $0 < i \leq size(P_h(s, d))$ and $0 < i' \leq k$.

In case of multiple selection options, $Hp_a(s,d) = Hp_b(s,d)$, the priority can be based on the shortest length, if $Lp_a(s,d) < Lp_b(s,d)$ then $p_a(s,d) \in P'_h(s,d)$ being $p_a(s,d), p_b(s,d) \in P_h(s,d)$.

Minimum Length Routing, ML:

This type of routing is interesting in order to minimize the fiber deployed, to support a given traffic demand, or to reduce the unavailability of paths due to cable cuts. The frequency of this type of failures is proportional to the length of the links involved.

The procedure of selecting k disjoint paths is the similar as for Minimum Hop Routing:

Definition 5.2. $p_{i'}(s,d)$ is a selected disjoint path in $P'_h(s,d)$ if $Lp_{i'}(s,d) \in min^k(Lp_i(s,d))$ and does not share any element with any of the other paths in $P'_h(s,d)$, for $0 < i \le size(P_h(s,d))$ and $0 < i' \le k$.

Traffic Balancing Routing, TB:

This type of routing is interesting, for example, in connection with congestion control problems. The routing itself presents an optimization problem. The set of paths between different node pairs are dependent on each other, and consequently, the complexity added to the problem is considerably increased. The main objective is to distribute traffic as evenly as possible across the network.

The goal of the optimization is to minimize variance between the traffic in the different links. In this case, the selection of k disjoint paths is defined as follows:

Definition 5.3. $p_{i'}(s,d)$ is a selected disjoint path in $P'_h(s,d)$ when $MIN(\triangle Ld)$ and does not share any element with any of the other paths in $P'_h(s,d)$. $\triangle Ld$ being $max(Ld_{l_{ij}} - Ld_{l_{i'j'}})\ \forall l_{ij}, l_{i'j'} \in S_L$ and Ld the link traffic load.

5.3. *Deployment Cost*

There are three main contributing elements regarding the economic costs for deploying optical transport networks: trenching, nodes, and fiber spans. These are used to define the following cost model, but obviously other elements could be included at a later stage. The model illustrates one of the possible methods to determine the economic costs of the deployment. Some of these concepts can be found in [21].

Let $EC_t = EC_l + EC_n$ be the total cost of deploying a network. EC_l is the cost of deploying the links and EC_n is the cost of deploying the nodes. Each link is characterized by its length, l_{ij}, and the traffic traversing it, Ld_{ij}. Basically, three cost constants can be defined to calculate EC_l: I_{trench}, I_{fix}, and I_{span}.

I_{trench} corresponds to the price for the trenching tasks per km and its value can significantly vary, depending on the region or landscape. I_{fix} corresponds to the fiber terminating equipment, and I_{span} is the cost related to each fiber span where ducts, fiber and amplifiers costs are included. The length of each span l_{span} can vary from 50 to 100 km.

14

Therefore, CW being the λ capacity and W being the number of $\lambda's$ per fiber, the number of fibers nf_{ij} for the link between i and j is defined in Eq. (5.2). Consequently, the economic costs for deploying a link, EC_l, is formally defined in Eq. (5.3).

$$nf_{ij} = \left\lceil \frac{Ld_{ij}}{CW \cdot W} \right\rceil \tag{5.2}$$

$$EC_l = \sum_{\forall\, i,j} I_{trench} \cdot l_{ij} + 2 \cdot nf_{ij} \cdot I_{fix} + \left(\left\lfloor \frac{TRF_{ij}}{l_{span}} \right\rfloor + 1 \right) \cdot I_{span} \tag{5.3}$$

Regarding the nodes, EC_n can be divided in two parts, the facility cost, I_{fal}, and the switching cost, I_{swch}, related to each switch size. The switch size is given by the number of incoming and outgoing fibers to each node and the number of wavelength per fiber. Usually standard switch sizes are given by $2^m X 2^m$ for $0 < m \leq 5$, and I_{swch} is not linearly proportional to m. Concluding, EC_n is formally defined as Eq. (5.4).

$$EC_n = N \cdot I_{fal} \sum I_{swch}(i) \,\forall\, i \,\in\, S_N \tag{5.4}$$

5.4. *Availability*

Availability is a convenient parameter for measuring the efficiency of the network infrastructure regarding failure support. Availability is defined in [4] as follows:

Availability is the probability of the system being found in the operating state at some time t in the future, given that the system started in the operating state at time t=0. Failures and down states occur, but maintenance or repair actions always return the system to an operating state.

In other terms, availability can be defined as the combination of reliability and restoration of a system [22]. This definition describes exactly the real goal of future networks in terms of failure support.

Availability in optical networks has been widely studied. For example, [23] presents an interesting availability review of Wavelength-Division Multiplexing (WDM) network components and systems. In terms of connection availability analysis, [24] compares the availability results between simulation and analytical environments for a shared protection scenario. Also, in [25] an interesting approach is followed to evaluate the availability in optical transport networks. It is measured in expected traffic losses when failures occur.

Availability is mathematically defined as the *working time/total time* ratio. Let $MTTF$ be the Mean Time To Fail of any element or system and MFT be the Mean Failure Time. Availability A is presented in Eq. (5.5), resulting in a numerical value of $0 \leq A \leq 1$.

$$A = \frac{MTTF}{MTTF + MFT} \tag{5.5}$$

The availability in relation to a $s - d$ pair connection is given by the availability calculation of sets of elements in serial and parallel. For each provided disjoint path to be available, all of its elements must be available. For a connection to be available, at least one path must be available.

Let A_{pj} be the availability of each of the provided k disjoint paths. m_j is the number of elements of A_{pj}, each of these characterized by an availability A_i. Eq. (5.6) presents the calculation of the availability for each path. The connection availability for a $s - d$ pair, A_{sd}, is presented in Eq. (5.7).

$$A_{pj} = \prod_{i=1}^{m_j} Ai \tag{5.6}$$

$$A_{sd} = 1 - \prod_{j=1}^{k}(1 - A_{pj}) \tag{5.7}$$

For this work, the availability figures only consider the physical hardware failure or cables cut. Extended models may include other factors such as power outages, human errors or blocking probabilities due to wavelength unavailability among others.

5.5. *Lifetime CO_2 releases*

Green House Gases (GHG) releases have been an important discussion topic, not only in scientific but also in political environments. The CO_2 contributions of IT systems releases have become more significant in the past few years [26]. Energy efficiency in terms of routing has been treated in studies such as [27], based on turning on and off the equipment when required. In terms of GHG emissions, [28] presents an exhaustive study regarding the deployment of Fiber-To-The-Home (FTTH) access network, considering many different aspects, going from machinery emissions to recycling processes.

In relation to this work, the CO_2 releases of a network along its lifetime (E_{NT}) can be quantified by adding the different types of contributions: Deployment E_D, Transmission E_T, Maintenance E_M, and Waste E_W, as defined by Eq. (5.8).

$$E_{NT} = E_D + E_T + E_M + E_W \tag{5.8}$$

The calculation of these parameters is complex due to the uncertainty in long term values of traffic, or emission factors of the different components. However, very briefly summarized as a generalization, the CO_2 contributions due to transmissions tasks, E_T, are inversely proportional to the connectivity factor α. E_D, E_M, and E_W are proportional to the length of the network. Since the network length usually increases along with the connectivity factor, it is not possible to minimize all the emission contributions. The higher the connectivity of the network is, the lower E_T and the higher E_D, E_M, and E_W become; and viceversa.

16

The balance among the different contributions and their specific relation to the connectivity factor and the network length may vary for specific scenarios, but these guidelines and patterns are usually followed. For a detailed description of this topic, it is recommended to read [29].

6. Solution search and Evaluation

This section presents the iterative process of problem solving that includes search algorithms and some of the most interesting objective functions. The overall goal is to simplify and improve the physical interconnection decision support procedure in optical transport networks. The combinatorial complexity of the problem requires a special attention on the reduction of the potential solutions to be considered. In addition, the proper definition of the objective function in relation to the specific requirements, might improve the outcome of the process. The following subsection deals in depth with these critical issues.

6.1. *Solution Search*

It is essential to work with "optimal" samples in order to extract conclusions from the analysis and evaluation of models and experiments. Specifically in this work, the search methods are optimization processes, characterized by the relevant objective function to the problem's nature.

There are many different options for decision making or solution search problems. The easiest idea is to solve the problem using *brute force* or *exhaustive search* algorithms. These are very simple, but unfortunately, the size of the problem makes their use inconvenient or even infeasible [30].

In this specific problem, the efficiency and versatility of heuristics qualify them as an interesting option. It is not the intention of this work to research in depth the different algorithms themselves. Instead, the purpose is to illustrate their application for network planning in order to create the proper environment for data collection and analysis.

The time frame when dealing with this kind of problems is essential. Experiments might require thousands of samples to be performed and these must be individually optimized; therefore, the processing time is limited. The absolute value of the time consumption is more related to the evaluation function complexity than to the search method itself. In fact, when dealing with heuristics to solve complex problems, the most time consuming tasks are the evaluation function subroutines. Theoretically, this set of combinatorial problems can be solved by Integer Linear Programing (ILP) techniques, but practically, heuristics are used due to time and size restriction.

Greedy algorithms are a priori, an interesting possibility since they are quite simple to implement, and are commonly used in some optimization problems. However, this type of algorithms provides local optima rather than global optima [30]. Instead, more complex algorithms such as Genetic Algorithms (GA), Simulated annealing (SA), or Tabu Search have succeeded when solving similar problems.

In [31] these are used to minimize the number of fibers deployed, and the number of wavelengths required to cover the traffic demands of a WDM network. In a similar way [21] deals with routing optimization, in order to efficiently dimension optical networks using SA. GA is used in [32], where the topology is optimized to minimize the number of optical cross-connect switches (OXCs), to serve a number of label switched routers (LSRs). However, the most of this kind of studies do not consider the physical deployment in their cost functions, or the physical location of nodes.

6.2. *Complexity analysis*

The next few paragraphs formally describe the complexity of the interconnection problem in order to understand its dimension. Mainly, there are two approaches to this type of problems: Topology related approach, and Link related approach, briefly introduced in Section 4. For further information about complexity concepts, it is recommended to read [30] .

Topology Related Approach:

For this approach, the proposal is to solve the problem using several well-known topologies, and select the option that suits the requirements best. The proposed topologies can be defined as *organized*, since they follow well defined patterns. These include regular, symmetric, or planar forming tessellation structures topologies. All of them have been studied in depth from a graph theory perspective.

Since this approach must be associated to an adjacency matrix, C^{Top}, only one topology can be optimized/evaluated at a time. This feature might be a constraint, but it considerably simplifies the problem. The solution of the interconnection problem following this approach can be formally described by:

Definition 6.1. Let $V = (\nu_1, \nu_2, \nu_3, ..., \nu_N)$ be a potential solution for the interconnection problem of the set of nodes $S_N = (N_1, N_2, N_3...N_N)$. V is a solution if, and only if, the three conditions in Eq. (6.1) are fulfilled.

$$\begin{aligned}
&\nu_i \leftarrow N_j, \ 0 < i, j \leq N \\
&N_j \ \in \ V, \ \forall \ 0 < j \leq N \\
&Size(S_N) = Size(V)
\end{aligned} \tag{6.1}$$

The relation between a solution vector V and the graph represented by an adjacency matrix C^{Top} is formally defined by:

$$\begin{aligned}
&if \ C_{ij}^{Top}(V) = 1 \ \Rightarrow Link \ between \ \nu_i \ and \ \nu_j \\
&\forall \ i, j \leq N
\end{aligned} \tag{6.2}$$

The most interesting property of this solution's format is that, solutions are always feasible as long as they contain all the elements of S_N only once. This leads to the statement in Th. 6.1, generically for any topology, and Th. 6.2 for a symmetric topology.

18

Theorem 6.1. V being a solution vector fulfilling Eq. (6.1), all the possible permutations of the elements of V lead to feasible solutions. Hence the, maximum number of different feasible solutions is $N!$.

Theorem 6.2. V being a solution vector fulfilling Eq. (6.1), all the possible permutations of the elements of V lead to feasible solutions. Moreover, if the topology is symmetric, the solution space is divided in sets of N equivalent solutions. Hence the maximum number of different feasible solutions is $N - 1!$.

Proof. Let G_0 be the resulting graph of assigning each N_i element to a ν_j, fulfilling Eq. (6.1), and following C^{Top}. Let G_m be the resulting graph of assigning each N_i element to a $\nu_{(j+m) \bmod N}$, fulfilling Eq. (6.1), following C^{Top}, and maintaining the same order for j as in G_0. If Top is symmetric, $G_0 = G_m$, $0 \leq m < N$. Therefore, each solution, G_0, has $N - 1$ exact replicas, concluding that the maximum number of different possible solutions is $N!/N = (N - 1)!$. $\qquad \square$

Following this procedure, if the problem is solved as optimization and evaluation of a finite number x of defined topologies, the maximum number of different feasible solutions $Max(Sol_T)$ is defined by Eq. (6.3). Note: Some of these might be duplicated.

$$Max(Sol_T) = x \cdot N! \tag{6.3}$$

Link Related Approach:

For this approach, the number of links L, or the adjacency matrix C^{Top}, or both are not a given input parameter. In case that L is fixed but not C^{Top}, the theoretical number of solutions is defined by:

Definition 6.2. Let a binary vector be $W = (\mu_1, \mu_2, \mu_3, ..., \mu_{\frac{N \cdot (N-1)}{2}})$. Each μ_i corresponds to a specific node pair, being 1 for an existing link between that pair and 0 for a hole (no existing link between them).

Assuming that W is a feasible solution for L links fulfilling Eq. (6.4), the maximum number of possible solutions based on the binomial coefficient is given by Eq. (6.5).

$$\sum_{i=1}^{\frac{N \cdot (N-1)}{2}} \mu_i = L \tag{6.4}$$

$$Max(Sol_{Lfixed}) = \binom{\frac{N \cdot (N-1)}{2}}{L} = \frac{\frac{N \cdot (N-1)}{2}!}{(\frac{N \cdot (N-1)}{2} - L)! \cdot L!} \tag{6.5}$$

In practice, some of these combinations are duplicated or become infeasible depending on the given constraints to interconnect the nodes, i.e. the solution is not a connected graph. Infeasible solution generation can be avoided by specific rules

for specific problem constraints, but it is not possible to generalize due to the high dependency on the specific problem conditions.

Furthermore, if L and C^{Top} are variable, the solution space is even larger, and consequently more complex to optimize. As an example, let the number of links be variable within a reasonable range, from a minimum 2-connected to a minimum 4-connected graphs[a].

Thus, the range of the possible number of links is $L = [N, 2N]$. Introducing L as a variable to Eq. (6.5), it is possible to conclude that the maximum number of possible solutions $Max(Sol_{L}var)$ is calculated as Eq. (6.6).

$$Max(Sol_{Lvar}) = \sum_{L=N}^{2N} \frac{\frac{N \cdot (N-1)}{2}!}{(\frac{N \cdot (N-1)}{2} - L)! \cdot L!} \tag{6.6}$$

As a conclusion, the difference between optimizing based on topologies in Eq. (6.3) and based on links position in Eq. (6.6) can be significant. Further studies regarding the comparison between the performances of these organized topologies vs. other randomly or irregularly connected topologies, is essential in order to define the limitations of the topology related approach.

6.3. *Search Algorithms*

In general physical interconnection problems are computationally complex, and the procedures to find a feasible solution go from exponential to factorial complexity. In such cases, stochastic search algorithms are the feasible choices. Only two of the most commonly used algorithms are presented in the context of the physical network interconnection problem to illustrate the idea. As commented above, it is not the scope of this study to go deep into the details of these algorithms. They are basically used as one of the tools for solving the interconnection problem, and only relevant information related to this topic will be specifically treated. However, relevant literature is given for each of the algorithms in case further information is required.

Genetic Algorithm (GA):

Genetic algorithm is one of the most popular to solve complex combinatorial problems. It was first introduced in [33], and has been revised several times. The basic idea of the algorithm is inspired by the natural selection process in biological evolution. GA mimics the evolution process by generating an initial random population (parents). Then, few of the parents are selected (natural selection) to produce a new population (children). The selection criteria is based on the principle of best-fit. The population that does not fulfill the fitting criteria is discarded.

For the implementation, an initial population set, potential solutions, are randomly generated. In the topology related approach, it can be represented as a vector

---

[a]This range is chosen since a 2-connected graph is the minimum reliable interconnection scheme and a 4-connected graph is the estimated real maximum due to the degree (4) of the road network.

20

V, containing the nodes. In the link related approach, the solution corresponds to vector W, containing the links information. Then, the vectors V from individual population are combined to generate the new vectors as population. The selection of which individuals should be used to generate the offspring, can be done by various selection schemes i.e. *Roulette Wheel* or *Tournament Selection*, see [10] for details. The criteria of what individuals to keep are assessed by an objective function. To generate the next generation, there are two main ways how those selected individuals are combined: Crossover and Mutation. More information about the selection and reproduction of individuals can be found in [34]. A detailed implementation example in relation to the interconnection problem can be found in [35].

Simulating Annealing, (SA):

Another of the most used methods is Simulated Annealing. The search method is inspired by the annealing process in metallurgy. This technique uses heating and gradual cooling of materials, to increase the size of its crystal and reduce the defects. When a material is heated, the atoms of the material move more freely with higher energies. But, when gradual cooling is applied, the atoms slow down with lower energies, and settle to better intended positions. The algorithm for optimization was first introduced in [36]. For the implementation based on the analogy, the algorithm generates random solutions encoded in vector V, similarly to GA. The current solution is then replaced by another random solution, with a probability that depends both on the difference between the corresponding objective functions, and the parameter temperature T. The value of T is gradually decreased in each of the iterations. When T is large, the probability of change in solutions is higher, and decreases until it reaches zero. A detailed implementation example regarding the interconnection problem can be found in [5].

Other Similar Techniques:

Both GA and SA, with their original implementations and their variants, are applied in many complex optimization problems. These variants can be tweaked to better fit for a particular problem application. The most notable are Evolutionary Programming, Tabu search, Quantum Annealing, Ant Colony optimization, and Harmony Search etc.

6.4. *Objective Function*

As mentioned above, the solution search and problem solving is characterized by an objective function, related to the problem's nature or experiment's goals. The objective function is the goal of the problem, for example to minimize the cost. The computing time is highly dependent on these objective functions. The iterative process of generation of new solutions is rather fast, but the evaluation process might require complex subroutines, i.e. k shortest paths selection. In large scale problems, parallel computing can be beneficial in terms of time consumption.

For certain experiments, it might not be necessary to deal with traffic and routing; this considerably simplifies the process. For others, dimensioning or capacity

allocation is unavoidable, and consequently, more complex tasks are required in the search process. Based on these guidelines, and in this specific case of optical network planning, there are two main types of objective functions in terms of evaluation processing time: *Routing dependent* and *independent*.

The complexity of the evaluation for both types is $O(N^2)$ on top of the combinatorial complexity of the physical interconnection problem itself. However, routing dependent problems are significantly more time consuming, since for each individual solution, evaluation, path selections, traffic distribution and other related tasks must be performed.

Some of the main objective functions are described in the following paragraphs:

Minimum Network Physical Length, $MIN(l_{nt})$

This routing independent optimization is very convenient in order to provide estimations about the economic cost or connection availability of the network, in a much shorter time than the following *Minimum Deployment Cost* and *Maximum Average Connection Availability* objective functions. Usually, in optical networks, trenching is significantly the most expensive task of the deployment stage.

The numerical values or the budget percentage of these trenching tasks, are directly related to the specific area where the network is deployed. The price fluctuation from country to country might be so high that it is not possible to generalize on this argument, but the minimization of the total length most likely involves a minimum, or near minimum, economic cost of deployment.

Minimum Deployment Cost, $MIN(EC_t)$

This optimization is routing dependent since the number of fibers per link, the number of amplifiers, and the switch sizes are dependent on the traffic distribution. As mentioned before, the resulting interconnection when minimizing the network's length and the economic cost is usually the same, keeping always in mind that the result is highly dependent on the trenching cost, I_{trench}.

Traditionally, this is the most popular optimization; however, this minimization might compromise the reliability of the network, if the constraints are not properly defined, i.e. minimum cut to avoid single points of failure.

Maximum Average Connection Availability, $MAX(\overline{A_{sd}})$

This objective function is routing dependent, and it leads to the most reliable network possible within the limitations provided by the constraints. Ideally, the best possible solution is a fully connected network, all-to-all physical links. But practically, constraints such as budget limitation and road network degree, affect the connectivity factor of the solution. A variance of this function is to maximize the *Minimum Connection Availability* for any $s-d$, optimizing in this case the worst possible case in the specific scenario.

Availability has not been widely used as an optimization objective function. One of the few examples in literature is [37], where availability is optimized when dealing with share path protection. However, the physical interconnection is a fixed input parameter and not the output result.

22

Minimum Lifetime CO_2 emission, $MIN(CO_2)$

The physical interconnection may have an influence on CO_2 releases in terms of network deployment and network operations. In this function, the energy efficiency factor can be implicitly included in the relation with the transmission emission, E_T, implying a routing dependent optimization. Energy consumption in optical networks is mainly related to the transmission of information and equipment maintenance, i.e. cooling systems. Ideally, this minimization results in a balance between the network length and the connectivity factor.

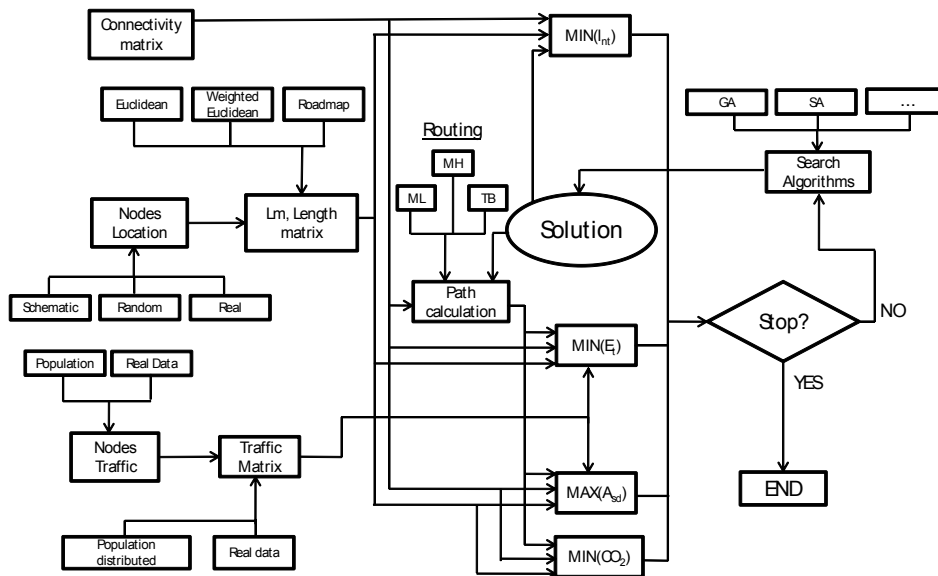Fig. 4 illustrates the flow process of the methodology described, to conclude with the modules' description.

Fig. 4.   Methodology Diagram

7. Use of the Methodology

This section describes examples of the possible uses of the methodology. Four experiments are briefly introduced and summarized, following the format: Problem statement, Description, Methods, and Conclusion. Some of the experiments are extracted from the authors' publications [38], [39] and [40], and these documents provide more detailed information about the procedure and the results. These experiments contributed to the definition of the proposed methodology. The presentation of these examples is kept simple, and only for illustration purposes; these are not the goal of

this specific work, and further reading is recommended for more detailed information.

7.1. *Genetic Algorithm Success Rate*

Problem statement: Determine the numerical values of the GA parameters to provide a probability of error, $P(e) < 0,01$ when solving an interconnection problem.

Description: The objective is to determine the value for the population size and number of iterations for GA, in order to provide an optimal result within confidence interval of 99% for any regular cyclic topology of $N = 20$ nodes. This type procedure is very convenient to validate the implemented algorithms.

Methods: This problem can be empirically solved by repeating the process several times. Using schematic nodes' location, the solution to the problem is known beforehand. The problem is solved for a regular cyclic topology, a Ring, 1000 times for each case. Each case is given a different population size, and the iteration when the optimum is achieved is recorded. The objective function used is *Minimum Network Physical Length* for computational time reasons. The results can be extrapolated to other regular cyclic topologies, since the complexity order is a function of N and not of L, as concluded in in Section 6, $Max(Sol_T) = (N-1)!$.

Conclusion: The number of required iterations to guarantee a $P(e) < 0,01$ is given by the 99th percentile of all the recorded number of iterations, to achieve the optimal solution in each case. Table 2 presents the average and 99th percentile of the required iterations to succeed in relation to the population size.

Table 2.   Average and 99th percentile of iterations to succeed

Population	N	2N	3N
Average	531.9	387.2	3264
99th Percentile	1528	1030	807

7.2. *Topology Length Prediction*

Problem statement: Determine the mathematical relation, if any, between the size of the area covered by the nodes to interconnect and the physical length of a network following a certain topology.

Description: The objective is to empirically find the patterns followed by different topologies, in relation to the size of the area covered by the nodes. In this way, their length can be estimated before any optimization is performed. As a result, the length for each topology should be characterized by approximation equations, as a function of the area covered and an estimated error. Detailed information about this and other similar experiments can be found in [38].

Methods: This experiment consists of two parts. Firstly, to determine the approximation function, several random nodes' locations interconnection problems are

24

solved for different topologies. In this case, the objective function is *Minimum Network Physical Length* for computational time reasons. In each case, a random number of nodes and coverage are given within a certain range. Then, all the samples are collected, and the approximation function is defined in relation to the size of the area and the number of nodes. Also, an estimated error must be provided.

Secondly, to verify the approximation, the same problem is solved several times with real geographical nodes' locations, while also varying the number of nodes. Finally, the result is compared with the approximation function value.

For this example, these approximation functions are calculated for a Ring and a Double Ring. Then they are applied to the real scenario of Nordjylland, in Denmark; one hundred times for each topology.

Conclusion: The length of the Ring implemented in Nordjylland (Denmark) was predicted with an average error of 6.8%, and a maximum error of 17%. The Double Ring was predicted with an average error of 8.8%, and a maximum error of 20%.

7.3. *Preventive Upgradeable Planning*

Problem statement: Determine the benefits and risks of preventive planning, using planar topologies.

Description: Certain organized topologies can be converted from one to the other, simply by adding or removing certain links. Planar topologies are very convenient in this aspect. Applying this concept, the network's infrastructure can be planned as optimized lower degree interconnection, and then new links are added to upgrade the network, or it can be optimized as a higher degree interconnection, and implemented in stages forming in each stage an organized topology. Both of these methods have risks and benefits. In the first case, the upgraded final topology is not optimized, and in the second case, if the final topology is not deployed for certain reasons, the middle topologies are not optimized. By comparing both options, it is possible to evaluate their risks and benefits. These concepts and experiments are taken from [15].

Methods: Three topologies are used for this experiment, Ring, Honeycomb, and Grid. From the Ring it is possible to form the Honeycomb by adding some links, and from the Honeycomb to form the Grid. The procedure consists of optimizing the interconnection of several scenarios, varying the number of nodes and the area covered by the nodes within a certain range. The objective function used is *Minimum Network Physical Length*. The same scenarios are optimized as the different topologies, and then the results are compared.

Conclusion: If the interconnection is optimized for a Honeycomb, the length of the Grid, when upgrading it, is 13 % longer on average than the optimal Grid. Instead, if the interconnection is optimized as a Grid, the length of the previous stage as a Honeycomb is 11% longer on average than the optimal Honeycomb. This implies that if the Grid is never deployed, the final network is not optimized.

7.4. *Fairness in Link Distribution*

Problem statement: Influence of equally distributing the links to interconnect a set of nodes forming regular topologies, and consequently, all the nodes have the same degree.

Description: The goal is to compare the deployment costs and availability of different physical network interconnections, for the same set of nodes and the same number of links. In this way, the effect of the distribution of the links to interconnect the nodes, can be quantified by economic and performance parameters. These experiments are extracted from [39] and [40]. Detailed information about the procedure is provided in these documents.

Methods: The experiment is performed using real nodes' locations taken from NSFNET, formed by 16 nodes and 24 links. Firstly, the relevant parameters are calculated for the interconnections, the costs and the availability. Secondly, three regular topologies are used to interconnect the same nodes. The result of connecting 16 nodes with 24 links, is a degree 3 and 3-connected regular topologies. These topologies are Double Rings and Chordal Rings. The objective functions used, are *Minimum Deployment Cost* and *Maximum Average Connection Availability*. In this way, it is possible to illustrate the cost implication of optimizing the availability, and the availability implications of optimizing costs. Furthermore, the comparison with the result of the traditional NSFNET interconnection illustrates the consequences of equally distributing the links in the network.

Conclusion: It is possible to equally distribute the links in the network under similar costs, improving the availability of the connections. Furthermore, as expected, the regular topologies are less affected by single and dual link failures.

8. Conclusion

The fast growth experienced in telecommunications in the past years, is leading toward high performance demanding information transport systems. The physical interconnection of networks plays a key role in reliability, availability, and QoS of transmissions. New design perspectives besides the "the best performance possible for the lowest budget", are required in order to improve the quality of the services provided. For this reason, it is important to define how to move forward in relation to this topic.

This paper describes a methodology in order to scientifically advance in the physical interconnection of optical transport networks. The methodology is composed of three main groups of modules: **Pre-calculated information** covers nodes location, links length, traffic, and interconnection schemes; **characteristic parameters** covers costs, availability, CO_2 emissions, and routing; **search and evaluation** cover objective functions, complexity analysis, and search methods.

The goal of this methodology is to define the techniques to deal with theories or hypothesis, to improve the current concepts when designing and deploying optical transport networks. It can be used for research purposes in many different aspects

26

of networks interconnection. One of the most important contributions is the formal definition of the procedure and methods to follow, considering a more global overview of the problem, other than only marginally solving parts of it.

For the first time, an overall methodology to cover many different aspects of the physical network interconnection has been proposed. This contribution can directly benefit the development in the direction of more reliable optical transport network and by extension, can contribute to the proper evolution of real long haul communication systems in the future.

Finally, some application examples are presented in order to illustrate the use of the different modules to deal with these problems. In addition, as part of its modularity, this methodology is left open for new ideas to be included as technology evolves. There are still many unsolved questions waiting to be covered.

9. Acknowledgment

References

1. R. Zhang-shen and N. Mckeown, "Designing a predictable internet backbone with valiant load-balancing," in *in IWQoS, 2005*, pp. 178–192, 2005.
2. K. Raza and M. Turner, *CCIE Professional Development Large-Scale IP Network Solutions*. Cisco Press, November 1999.
3. M. To and P. Neusy, "Unavailability analysis of long-haul networks," *IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS*, vol. 12, January 1994.
4. W. D. Grover, *Mesh-Based Survivable Networks, Options and Strategies for Optical, MPLS, SONET and ATM Network*. Prentice Hall PTR, 1st edition ed., August 2003.
5. T. Riaz, *SQoS based Planning for Network Infrastructures (pages 125-137)*. Department of Electronic Systems, Aalborg University, 2008.
6. O. B. Madsen, T. P. Knudsen, and J. M. Pedersen, "Sqos as the base for next generation global infrastructure," in *Proceedings of IT&T 2003, Information Technology and Telecommunications Annual Conference 2003*, pp. 127–136, 2003.
7. M. Dehmer, "A novel method for measuring the structural information content of networks," *Cybern. Syst.*, vol. 39, pp. 825–842, November 2008.
8. M. Dehmer, S. Borgert, and F. Emmert-Streib, "Network classes and graph complexity measures," in *Proceedings of the 2008 First International Conference on Complexity and Intelligence of the Artificial and Natural Complex Systems. Medical Applications of the Complex Systems. Biomedical Computing*, (Washington, DC, USA), pp. 77–84, IEEE Computer Society, 2008.
9. R. Huang, J. Ma, and D. F. Hsu, "A genetic algorithm for optimal 3-connected telecommunication network designs," in *Proceedings of the 1997 International Symposium on Parallel Architectures, Algorithms and Networks*, ISPAN '97, (Washington, DC, USA), pp. 344–350, IEEE Computer Society, 1997.
10. A. Vivas, L. de Errico, C. H. Barbosa, A. L. L. Aquino, and L. P. de Assis, "Analysis of selection and crossover methods used by genetic algorithm-based heuristic to solve the

lsp allocation problem in mpls networks under capacity constraints," in *International Conference on Engineering Optimization, EngOpt 2008*, pp. 1–4, 2008.

11. S. Baroni and P. Bayvel, "Wavelength requirements in arbitrarily connected wavelength-routed optical networks," *Lightwave Technology, Journal of*, vol. 15, no. 2, pp. 242 –251, 1997.

12. A. Dacomo, S. D. Patre, G. Maier, A. Pattavina, and M. Martinelli, "Design of static resilient wdm mesh networks with multiple heuristic criteria," in *IN PROCEEDINGS OF IEEE INFOCOM2002*, pp. 1793–1802, 2002.

13. R. Singel, "Fiber optic cable cuts isolate millions from internet, future cuts likely," January 2008.

14. R. Diestel, *Graph Theory (Graduate Texts in Mathematics)*. Springer, August 2005.

15. J. M. Gutierrez, T. Riaz, J. M. Pedersen, and O. B. Madsen, "Upgradeability and predictability analysis for mesh topologies in optical distribution networks," in *Wireless and Optical Communications Networks, 2009. WOCN '09. IFIP International Conference on*, pp. 1 –5, April 2009.

16. J. Fernandez, P. Fernandez, and B. Pelegrin, "Estimating actual distances by norm functions: a comparison between the lk,p, theta;-norm and the lb1, b2, theta;-norm and a study about the selection of the data set," *Comput. Oper. Res.*, vol. 29, no. 6, pp. 609–623, 2002.

17. R. Sinnott, "Virtues of the haversine," *Sky and Telescope*, vol. 68, no. 2, p. 159, 1984.

18. J. M. Pedersen, *Structural Quality of Service in Large-Scale Network*. PhD thesis, Aalborg University, April 2005.

19. T. S. El-Bawab, M. J. O'Mahony, and A. P. Jayasumana, "A european multiwavelength optical network," in *ISCC '97: Proceedings of the 2nd IEEE Symposium on Computers and Communications (ISCC '97)*, p. 568, IEEE Computer Society, 1997.

20. G. N. Rouskas and H. G. Perros, "A tutorial on optical networks," in *Advanced lectures in Networking*, pp. 155–193, 2002.

21. B. V. Caenegem, W. V. Parys, F. D. Turck, and P. Demeester, "Dimensioning of survivable wdm networks," *Selected Areas in Communications, IEEE Journal on*, vol. 16, pp. 1146 –1157, September 1998.

22. B. Warfield, "Design of availability, reliability, and restoration for a telecommunications access network guided by customer demand for quality," in *IN PROCEEDINGS 2010 Australasian Telecommunication Networks and Applications Conference*, pp. 125–130, 2010.

23. I. Rados, "Availability analysis and comparison of different wdm systems," *Journal of Telecommunications and Information Technology*, vol. 1, pp. 114–119, 2007.

24. L. Zhou, M. Held, and U. Sennhauser, "Connection availability analysis of shared backup path-protected mesh networks," *J. Lightwave Technol.*, vol. 25, pp. 1111–1119, May 2007.

25. G. Booker, A. Sprintson, E. Zechman, C. Singh, and S. Guikema, "Efficient traffic loss evaluation for transport backbone networks," *Comput. Netw.*, vol. 54, pp. 1683–1691, July 2010.

26. C. Panarello, A. Lombardo, G. Schembra, L. Chiaraviglio, and M. Mellia, "Energy saving and network performance: a trade-off approach," in *Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking*, e-Energy '10,

28

(New York, NY, USA), pp. 41–50, ACM, 2010.

27. L. Chiaraviglio, M. Mellia, and F. Neri, "Reducing power consumption in backbone networks," in *In IEEE International Conference on Communications (ICC09)*, 2009.

28. E. for FTTH Council Europe, "Ftth solutions for a sustainable development," February 2008.

29. J. Gutierrez, T. Riaz, J. M. Pedersen, A. Patel, and O. B. Madsen, "Analytical model based on green criteria for optical backbone network interconnection," *Computer Standards & Interfaces*, vol. 33, no. 6, pp. 574 – 586, 2011.

30. H. S. Wilf, "Algorithms and complexity," 1994.

31. H. Liu and F. A. Tobagi, "Physical topology design for all-optical networks," in *BROAD-NETS*, 2006.

32. Y. Xin, G. N. Rouskas, H. G. Perros, S. Member, and S. Member, "On the physical and logical topology design of large-scale optical networks," *Journal of Lightwave Technology*, vol. 21, pp. 904–915, 2003.

33. J. H. Holland, "Genetic algorithms and the optimal allocation of trials," *SIAM J. Comput.*, vol. 2, no. 2, pp. 88–105, 1973.

34. B. Al-Bassam, A. Alheraish, and S. H. Bakry, "A tutorial on using genetic algorithms for the design of network topology," *International Journal of Network Management*, vol. 16, pp. 253 – 262, July 2006.

35. J. M. Gutierrez, M. Imine, and O. B. Madsen, "Network planning using ga for regular topologies," in *Proceedings of IEEE International Conference on Communications, ICC 2008, Beijing, China, 19-23 May 2008*, pp. 5258–5262, 2008.

36. S. Kirkpatrick, D. G. Jr., and M. P. Vecchi, "Optimization by simmulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

37. M. Tornatore, G. A. Maier, A. Pattavina, M. Villa, A. Righetti, R. Clemente, and M. Martinelli, "Availability optimization of static path-protected wdm networks," in *Optical Fiber Communication Conference*, Optical Society of America, 2003.

38. J. M. Gutierrez, T. Riaz, J. M. Pedersen, and O. B. Madsen, "Influence of schematic network topology models over real physical networks," in *ICACT'09: Proceedings of the 11th international conference on Advanced Communication Technology*, pp. 599–604, IEEE Press, 2009.

39. J. M. Gutierrez, K. Katrinis, K. Georgakilas, A. Tzanakaki, and O. B. Madsen, "Increasing the cost-constrained availability of wdm networks with degree-3 structured topologies," in *Transparent Optical Networks (ICTON), 2010 12th International Conference on*, pp. 1–4, 2010.

40. J. M. Gutierrez, K. Georgakilas, T. Riaz, K. Katrinis, J. M. Pedersen, A. Tzanakaki, and O. B. Madsen, "Single and dual physical link failures stability effect on degree three wdm networks," in *The 16th Asia-Pacific Conference on Communications Proceedings, APPC 2010*, 2010.