Genetic Algorithm for the Topological Design of Survivable Optical Transport Networks

Rui Manuel Morais, Claunir Pavan, Armando Nolasco Pinto, and Cristina Requejo

Abstract—We develop a genetic algorithm for the topological design of survivable optical transport networks with minimum capital expenditure. Using the developed genetic algorithm we can obtain near-optimal topologies in a short time. The quality of the obtained solutions is assessed using an integer linear programming model. Two initial population generators, two selection methods, two crossover operators, and two population sizes are analyzed. Computational results obtained using real telecommunications networks show that by using an initial population that resembles real optical transport networks a good convergence is achieved.

Index Terms—Genetic algorithm; Survivable network design; Optical networks; Dedicated protection; CAPEX.

I. INTRODUCTION

The proliferation of broadband access networks is continuously demanding flexible, scalable, and reliable transport networks with minimum capital expenditures (CAPEX) and operational expenditures (OPEX). The CAPEX of an optical network can be divided into costs for bandwidth management (costs with nodes) and costs for signal transmission (costs with links) [1]. The node location is one of the first pieces of information that the network designer has, corresponding to the location of the central offices where the traffic is added and dropped. The first stage of the overall network design process is the topological design; at this stage the connections between the nodes are established. The network topological design should guarantee a reliable network, and this depends on which links are going to be implemented [2]. The traffic to be transported by the network is hard to forecast and is continuously changing [3]. In practice, several traffic scenarios are defined and evaluated, then the lowest cost network that will remain feasible for the majority of the scenarios is implemented [3]. Therefore, the utilization of methods to quickly design

physical topologies ensuring the routing of the required traffic and guaranteeing the network survivability at minimum cost is crucial. In this work, we consider that the network should be survivable against any single link failure. Thus, the underlying topology is a 2-connected graph [4]. The topological design of minimum cost 2-connected graphs, not allowing the use of parallel edges, is strongly NP-hard [2,5]; thus integer linear programming (ILP) models only lead to optimal solutions, within reasonable time and computational effort, for small networks. Consequently, heuristics are commonly used to search for near-optimal solutions.

The survivable topological design problem has attracted the attention of many researchers [2,6–10]. A survey on ILP models, decomposition methods, and heuristics for the topological design of survivable networks can be found in [2,6]. In [8–10] ILP models and heuristics are analyzed to minimize the total number of links. Optimization methods to minimize the CAPEX of an optical WDM network can be found in [11–14]. In [11], ILP models are presented to obtain the least-cost network in terms of CAPEX, for various network architectures, and in [12], deterministic and greedy heuristics are presented to optimize the lightpath routing with and without protection. A tool to jointly minimize the CAPEX and the OPEX is presented in [13]. In [14–16] genetic algorithms are used in the design of telecommunication networks. In [14], a genetic algorithm is presented to route and dimension dynamic optical networks, without considering the network survivability. In [15], a genetic algorithm is used to design survivable networks. It is assumed that the topology is given. The genetic algorithm presented in [16] minimizes the CAPEX of an all-optical network, considering physical impairments. To contour the NP-hardness nature of the problem, the physical topology and the traffic model are known in advance in [11–15]. In [1,17] simplified models are developed to estimate the optical network CAPEX without the knowledge of the network topology.

In this work, we address the problem of jointly designing the physical topology, ensuring survivability, and minimizing the network CAPEX of an opaque optical transport network with a transparent switch [18]. In order to deal with this problem we propose a genetic algorithm. As the convergence of the genetic algorithm depends on the used genetic operators, we analyze their impact on the quality of the obtained solutions. Two initial population generators, two selection methods, two crossover operators, and two population sizes are compared within the genetic algorithm. The

performance of the proposed heuristic is evaluated using an ILP model. We use a simplified cost model to calculate the CAPEX of an optical network to obtain exact results that can be compared with the heuristic solutions. The computational results are obtained using the node location of nine real telecommunications networks.

This paper is organized as follows: the problem is formulated in Section II. In Section III, we present the genetic algorithm and the ILP model is presented in Section IV. Computational results obtained with the genetic algorithm and with the ILP model are reported in Section V. Finally, in Section VI we draw the main conclusions.

II. PROBLEM FORMULATION

A. Network Architecture

In this work, we consider a multilayered network, with an electrical and an optical layer. The electrical layer is responsible for the traffic grooming and protection. We focus on path-dedicated protection, where a disjoint backup path is used to protect each optical channel [19]. Thus, the network is survivable against failure in any optical channel or in any single link. The optical layer has the function of multiplexing, transmitting, and switching wavelength signals to establish end-to-end paths [18].

An optical transport network can be seen as a set of nodes connected by bidirectional links. We assume that the traffic is bidirectional and follows the same path in both directions. Figure 1 presents the node architecture considered. We assume that the network nodes are equipped equally with an electrical cross connect (EXC) and an optical cross connect (OXC). The traffic enters into the transport network via the EXC, with different granularities. It is subsequently groomed into the fundamental units of capacity used in the WDM transmission systems and sent to the OXC. To ensure survivability, each groomed signal is routed through two link-disjoint paths. The optical signal traverses two or more OXCs in the path from the source to the destination node. In the intermediary nodes, the optical signal is bypassed at the optical layer (see dashed line in Fig. 1). In the terminal node it is switched to the EXC (see solid line in Fig. 1). In this way, the EXC only processes the local traffic. However, for regeneration purposes, we assume that optical signals suffer an optical–electrical–optical (OEO) conversion at every node [18,20,21].

The transponders mark the end points of wavelength signals and comprise a short-reach interface and a long-reach interface. At the node side the transponder has a short-reach interface, whereas at the line side it has a long-reach interface to send the optical signal along the transmission system. The transponders also have the function of wavelength assignment. The transmission system is composed of WDM terminals, optical amplifiers, and the optical fiber; see Fig. 2. The WDM terminal is capable of multiplexing/demultiplexing multiple wavelength signals into/from a single optical fiber. Besides this, the WDM terminal also has a booster amplifier and a preamplifier. Optical line amplifiers are installed along the transmission system to amplify the optical signal.

We assume that the node location and the traffic are inputs of the model. We are also assuming that the EXC and the OXC switching matrices are able to process all the required traffic. Therefore, the number of EXCs, OXCs, tributary ports, and ports between the EXC and the OXC are fixed.

B. Mathematical Formulation

The network has to support a given traffic, corresponding to a set of demands between nodes. The demand between the origin node, o, and the destination node, d, is denoted by (o,d) and the set of all demands by $D = \{(o,d) : o,d \in V\}$, where $V = \{1, \ldots, N\}$, is the set of nodes. The number of optical channels needed in each transmission system to support the demand (o,d) is denoted by B_{od}.

The network topology corresponds to a set of links connecting pairs of nodes. In each link, several transmission systems can be installed. The cost of all transmission systems in the link $\{i,j\}$ without transponders, F_{ij}, is the cost with its deployment plus the cost with the WDM terminals, optical amplifiers, and optical fiber. The cost of deploying a transmission system depends on several parameters such as geographic factors and method used [22]. The cost with the optical fiber is proportional to the length C_{ij} between the nodes. The number of optical amplifiers depends on C_{ij} and on the *span*, which is defined as the maximum allowed distance between consecutive optical amplifiers. Moreover, two WDM terminals are needed per transmission system, one in each end. Therefore, F_{ij} is given by

$$F_{ij} = \left(c_{term} + \left\lceil \frac{C_{ij}}{span} - 1 \right\rceil c_{oa} + C_{ij} c_f + M_{ij} \right) X_{ij}, \quad (1)$$

where c_{term} is the cost of two WDM terminals, c_{oa} is the cost of a bidirectional optical amplifier, $[(C_{ij}/span) - 1]$ is the number of optical amplifiers, c_f is the cost of the optical fiber per kilometer, and M_{ij} is the cost with right-of-way privi-
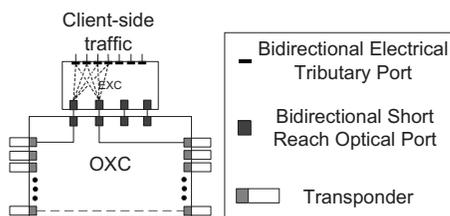

Fig. 1.   Node architecture: bidirectional electrical tributary ports, bidirectional short-reach optical ports, electrical cross connect (EXC), and optical cross connect (OXC).
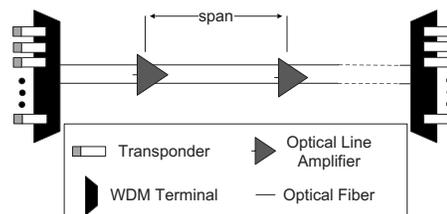

Fig. 2.   Transmission system architecture: transponders, WDM terminal, optical line amplifiers, and optical fiber.

leges and/or method used to deploy a transmission system between the nodes i and j. The integer nonnegative variable X_{ij} indicates the number of transmission systems installed between the nodes i and j. The number of transmission systems installed, X_{ij}, is dependent on the maximum number of optical channels that a WDM terminal can multiplex/demultiplex into/from a single optical fiber and on the number of optical channels that traverse the link $\{i,j\}$, L_{ij}. The value of L_{ij} is calculated by adding the number of optical channels, B_{od}, for all demands that traverse the link $\{i,j\}$. Thus,

$$L_{ij} = \sum_{(o,d) \in D} B_{od} Z_{ij}^{od}, \tag{2}$$

where Z_{ij}^{od} is a binary variable that indicates whether the demand (o,d) is routed through the link $\{i,j\}$. Therefore, X_{ij} is obtained by

$$X_{ij} = \left\lceil \frac{L_{ij}}{K_{ij}} \right\rceil, \tag{3}$$

where K_{ij} is the capacity of each transmission system in the number of optical channels. The cost with transponders, O_{ij}, corresponds to two transponders per optical channel that traverses the link. Given that c_t is the cost of a pair of transponders, O_{ij} can be calculated by

$$O_{ij} = c_t L_{ij}. \tag{4}$$

Therefore, the transmission cost, T_c, is given by the sum of the cost F_{ij} plus O_{ij} for all links,

$$T_c = \sum_{\{i,j\} \in E} (F_{ij} + O_{ij}), \tag{5}$$

where $E = \{\{i,j\}: i,j \in V, i < j\}$, is the set of edges that represents all possible bidirectional links.

The main goal of this work is to search for the physical topology that minimizes T_c. The topology and the paths of each demand can be obtained by the values of the variables X_{ij} and Z_{ij}^{od}, respectively. The topological design problem is characterized by being hard in complexity, time consumption, and memory requirements. Consequently, exact solutions within reasonable time can only be obtained for small networks. As optical transport networks can have more than 100 nodes, ILP models are prohibitive for larger networks. In the next sections, we present a heuristic approach and an ILP model to obtain solutions that minimize expression (5) with survivability requirements. The ILP model is used to evaluate the quality of the solutions obtained using the genetic algorithm in terms of accuracy and time consumption, for small networks.

III. GENETIC ALGORITHM

A genetic algorithm is a heuristic based on the theory of natural evolution and has the following steps: generation of an initial population, encoding, evaluation, selection, crossover, mutation, and decoding [23]. A set of initial feasible solutions of the problem (individuals) is generated, forming the initial population. Afterwards, the genetic algorithm modifies this population repeatedly. A pair of individuals is chosen under selection rules and the individuals are com-

bined under crossover rules, giving rise to another pair of individuals (offsprings). To increase the population diversity, mutations can also be applied.

The convergence of a genetic algorithm is dependent on the used operators. In this work, we compare two initial population generators, two selection methods, two crossover operators, and two sizes of the population within the genetic algorithm.

A. Initial Population

A feasible problem solution is a network topology with at least two link-disjoint paths between any pair of nodes. In this contribution we use two topology generators to create the initial population set. One generates completely random topologies, based on [17]. The other generates topologies that preserve the main characteristics of real telecommunication networks, based on [24].

The random topology generator starts by designing a ring topology connecting all nodes of the network, thus guaranteeing that all initial solutions are feasible [17]. The ring topology is randomly generated for each individual of the initial population. Afterwards, t links are added to the ring topology, connecting t pairs of randomly selected nodes. The number t of additional links is randomly generated and ranges from 0 to $(N^2 - 3N)/2$, i.e., from a ring to a full mesh network.

In the topology generator presented in [24], the nodes are placed into a plane that is divided into regions. Then, if there are two nodes located in the same region they will be directly connected; if there are three or more nodes in the same region they will be connected as a closed cycle. After this initial procedure, the network survivability is ensured by connecting each region to two other regions. Afterwards, a random number of additional links are added to the network. The probability $P(i,j)$ of a link existing between the node i and the node j is given by the Waxman link probability [25],

$$P(i,j) = \beta \exp \frac{-C_{ij}}{\alpha L}, \tag{6}$$

where C_{ij} is the distance, in kilometers, between the node i and j; L is the maximum distance between any two nodes in the network; and following [24], α and β are both assumed to be 0.4.

B. Encoding and Decoding

The encoding corresponds to the creation of a genetic code that uniquely represents a solution. To encode the solutions, we used the concatenation of the rows of the upper triangular matrix of the adjacency matrix. The adjacency matrix is an $N \times N$ matrix in which an element in position i,j is 1 if node i is directly connected to node j, and 0 otherwise. As the network links are bidirectional, the adjacency matrix is symmetric. As an example, consider the network topology presented in Fig. 3. The adjacency matrix of the solution presented in Fig. 3 is
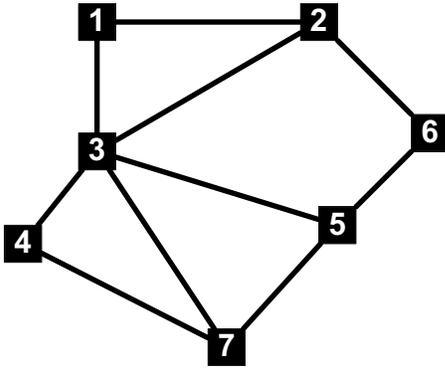
Fig. 3.   Network topology with seven nodes.

$$[g] = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}, \qquad (7)$$

and the respective genetic code is

1 1 0 0 0 0|1 0 0 1 0|1 1 0 1|0 0 1|1 1|0.

Given the genetic code of the solutions, the decoding is the inverse operation.

C. Evaluation

The evaluation consists of determining the transmission cost of each topology. To calculate the network cost we determine two paths for each demand (a working and a backup path) and use them to calculate the number of optical channels needed in each link. We assume that the working path is the shortest path (in number of hops) and the backup path is the second-shortest path. The working path is determined using the Dijkstra algorithm. To obtain a second link-disjoint shortest path, the links in the working path are overweighted, and the backup path is determined using the Dijkstra algorithm a second time. If two or more shortest paths exist, we use the first path encountered by the Dijkstra algorithm. Note that the feasibility of the solutions can only be guaranteed at the initial population. Later on, if the two disjoint paths cannot be obtained, the solution is removed from the population. The two link-disjoint paths determined using the Dijkstra algorithm will be used to give values to the variables Z_{ij}^{od}. After all the demands have been routed, O_{ij} is obtained using Eq. (4) and X_{ij} using Eq. (3). The cost F_{ij} is obtained using Eq. (1) and T_c is calculated using expression (5).

The use of the shortest paths in number of hops to route the demands is an approximation used in the genetic algorithm. This approximation is assessed in Section V using the ILP model presented in Section IV.

D. Selection

In the selection phase, pairs of individuals are chosen for crossover. Usually, individuals are selected based on their fitness, i.e., the cost of the respective solution, emphasizing the fitter individuals expecting that their offspring will have higher fitness. However, a strong selection can reduce the diversity of the population, leading to suboptimal solutions; contrariwise a weak selection can result in slow evolution [26]. We have implemented two different selection methods that differ in the selection pressure: the roulette wheel and the tournament method [23].

In the roulette wheel method, after all solutions are evaluated, the total generation cost is determined by adding all the solution costs. The next step is the calculation of the solution fitness. The fitness of each solution is the difference between the total generation cost and the cost of the solution. In this way, the solutions with lower cost will have higher fitness than the solutions with higher cost. Finally, the selection probability is calculated by the ratio between the solution fitness and the sum of the fitness of all individuals. Therefore, solutions with lower cost have greater probability of being selected for crossover.

In the tournament method, four individuals are randomly selected from the population and grouped two by two. Afterwards, two numbers, r_1 and r_2, ranging between 0 and 1 are randomly generated. If $r_1 < 0.75$, the solution with the smaller cost from the first group is selected for crossover; otherwise the less-fit individual is selected. The same occurs for the second group.

E. Crossover and Mutation

In the crossover operation, pairs of individuals previously selected are combined, giving rise to another pair of new individuals. Two crossover operators are implemented: the single point crossover and the uniform crossover [23].

In the single point crossover, a border between two elements of the genetic code is randomly selected. The two left sides of progenitor 1 and progenitor 2 are copied to offspring 1 and offspring 2, respectively. The right sides of each code shall be exchanged, i.e., the right side of progenitor 1 is copied to offspring 2 and the right side of progenitor 2 is copied to offspring 1. An example is displayed in Table I. In this example, the border is placed between the fourth and the fifth element of the genetic code for the two progenitors.

In the uniform crossover, a mask is randomly generated. If the crossover mask bit i is 1, the offspring 1 receives the bit i from progenitor 1 and offspring 2 receives the bit i from progenitor 2. If the mask bit i is 0, offspring 1 inherits the

TABLE I
EXAMPLE OF THE SINGLE-POINT CROSSOVER

Individual	Genetic Code					
Progenitor 1	1	1	1	1	0	1
Progenitor 2	1	0	1	1	1	1
Offspring 1	1	1	1	1	1	1
Offspring 2	1	0	1	1	0	1

bit i from progenitor 2 and offspring 2 inherits it from progenitor 1. The example in Table II illustrates this process.

The mutation operation consists of a simple exchange of 0's to 1's, or vice versa, at random locations of the genetic code, for a randomly selected number of individuals. This operation has the goal of increasing the diversity of the population.

After the individuals are evaluated, selected, and reproduced, the next generation is created. The selection of the individuals to form the next generation is made from the present generation and the generated offspring. We consider that a maximum of 20% of individuals are selected from the present generation, the remaining 80% are offspring, to make available for crossover the maximum number of links as possible.

IV. INTEGER LINEAR PROGRAMMING MODEL

In this section, we present an ILP model to minimize the cost of a survivable optical transport network [6,8–10]. In order to formulate the flow conservation constraints the binary variable Z_{ij}^{od} is divided into two variables Y_{ij}^{od} and Y_{ji}^{od}. The variable Y_{ij}^{od} indicates that the demand (o,d) is routed through the link $\{i,j\}$ in the direction from i to j. The variable Y_{ji}^{od} indicates that (o,d) is routed through the link $\{i,j\}$ in the direction from j to i. The ILP model is the following:

$$\text{minimize } T_c = \sum_{\{i,j\} \in E} (F_{ij} + O_{ij})$$

subject to

$$\sum_{j \in V \setminus \{o\}} Y_{ij}^{od} - \sum_{j \in V \setminus \{d\}} Y_{ji}^{od} = \begin{cases} 2, & i = o \\ 0, & i \neq o,d \\ -2, & i = d \end{cases}$$

$$\forall (o,d) \in D, \quad \forall i \in V, \tag{8}$$

$$\sum_{(o,d) \in D} B_{od}(Y_{ij}^{od} + Y_{ji}^{od}) \leq K_{ij} X_{ij} \quad \forall \{i,j\} \in E, \tag{9}$$

$$X_{ij} \in \mathbb{N}_0 \quad \forall \{i,j\} \in E, \tag{10}$$

$$Y_{ij}^{od}, Y_{ji}^{od} \in \{0,1\} \quad \forall (o,d) \in D \quad \forall \{i,j\} \in E. \tag{11}$$

The objective function, to be minimized, is expression (5). Constraints (8) are the usual flow conservation constraints and ensure that, for each (o,d) pair, we route two units of

TABLE II
EXAMPLE OF THE UNIFORM CROSSOVER

Individual	Genetic Code					
Progenitor 1	1	**1**	1	1	**0**	1
Progenitor 2	1	**0**	1	1	**1**	1
Mask	0	**1**	1	0	**0**	1
Offspring 1	1	**1**	1	1	**1**	1
Offspring 2	1	**0**	1	1	**0**	1

flow from node o to node d. These constraints together with constraints (9) guarantee the connectivity between all pairs of nodes. Constraint set (9) connects the sets of variables, guaranteeing that the total number of optical channels that crosses the link $\{i,j\}$, in both directions, does not exceed the maximum capacity, K_{ij}, of the number of installed transmission systems, X_{ij}. Constraint set (10) defines the variables X_{ij} as nonnegative integer variables, allowing the installation of more than one pair of transmission systems in each link. The disjointness of the two flows, to ensure survivability, is enforced by constraints (11). As the variables Y_{ij}^{od} and Y_{ji}^{od} are binary, the two flows cannot traverse the same edges. Hence, the existence of two link-disjoint paths for each origin destination pair is guaranteed.

V. RESULTS

In this section, the computational results obtained using the genetic algorithm and the ILP model are reported. The genetic algorithm is implemented in C++. Using the genetic algorithm, feasible solutions corresponding to upper bounds for the optimal value are obtained. The ILP model is used to obtain lower bounds and is solved using the branch and bound method from the commercial optimization software Xpress IVE 1.18. The results are obtained using a PC Intel Core 2 at 1.83 GHz and 1 GB RAM. The halting criteria used for the ILP model is the obtention of the optimal solution or 10 hours of processing time. In the genetic algorithm we performed 100 iterations, which required less than 4 minutes for the largest network. We observed marginal improvements in the solutions obtained when increasing this number of iterations.

To evaluate the quality of the obtained solutions the gap between the upper, b_u, and the lower, b_l, bound is calculated as follows:

$$gap = \frac{100(b_u - b_l)}{b_u}, \tag{12}$$

where b_u is obtained using the genetic algorithm and b_l using the ILP model.

The computational results are obtained for the node location of nine real telecommunications networks. We assume that all links can be implemented and that the cost of deploying a transmission system is independent of the link, i.e., $M_{ij}=M$. For a matter of simplicity we consider $M=0$. The maximum number of optical channels supported by each transmission system is 40, i.e., $K_{ij}=K=40$. We also consider that the maximum distance between optical amplifiers is 80 km, i.e., $span=80$. The cost with the equipment in normalized monetary units (m.u.) is presented in Table III.

To assess the impact of the initial population, selection method, crossover operator, and number of individuals in the population, we perform five runs for each combination:
- 100 individuals, roulette wheel selection, single-point crossover;
- 100 individuals, tournament selection, single-point crossover;
- 100 individuals, roulette wheel selection, uniform crossover;

Equipment	Notation	Cost (m.u.)	Quantity
Optical Fiber	c_f	0.80	per km
Optical Amplifier	c_{oa}	1.92	per fiber and per span
WDM Terminal	c_{term}	8.34	per fiber
Transponder	c_t	0.66	per fiber per channel

- 100 individuals, tournament selection, uniform crossover;
- 500 individuals, roulette wheel selection, single-point crossover;
- 500 individuals, tournament selection, single-point crossover;
- 500 individuals, roulette wheel selection, uniform crossover;
- 500 individuals, tournament selection, uniform crossover.

A. Impact of the Initial Population

We start by assessing and comparing the quality of the obtained solutions when using different initial population generators. The presented results are obtained using a uniform demand matrix. The initial populations are randomly generated [17] or following [24], as described in Subsection III.A. The number of nodes, the number of regions, and the number of nodes placed in each region are presented in Table IV for all considered networks.

Figure 4 shows the evolution of the gap for the best solution obtained, among all combinations, for initial populations generated using the random topology generator [Fig. 4(a)] and using the topology generator presented in [24] [Fig. 4(b)].

As can be seen, for initial populations generated using the random topology generator, the optimal solution is obtained for the three smallest networks. With the increase in the number of nodes the gap also increases reaching almost 15%; see Fig. 4(a).

For initial populations generated using the topology generator presented in [24] [see Fig. 4(b)], the genetic algorithm also obtains the optimal solution for networks with less than 12 nodes. The exception is the CESNET network, in which

TABLE IV
REAL-WORLD REFERENCE NETWORKS [24]

Network	Nodes	Regions	Nodes per Region
VIA	9	2	5-4
RNP	10	4	0-8-1-1
vBNS	12	3	3-4-5
CESNET	12	3	4-7-1
ITALY	14	2	12-2
NFSNET	14	2	7-7
AUSTRIA	15	3	3-4-8
GERMANY	17	4	8-2-5-2
SPAIN	17	4	8-2-7-0

the best solution obtained has a gap of 4.1%. Increasing the number of nodes, the gap also increases. However, the solutions obtained within 100 iterations always have gaps smaller than 10%; see Fig. 4(b).

Considering the best solution obtained for all networks and combinations, only for the networks in which the optimal solution is obtained, an initial population randomly generated obtains a solution as good as the one obtained using [24]. In all the other networks the solutions obtained using the topology generator presented in [24] have smaller cost. The improvements range between 1% and 10%. One reason for this is that in the random topology generator all the links have the same probability to be chosen. Contrariwise, in the topology generator presented in [24] longer links have smaller probability than shorter ones. We also used a ring-based random topology generator in which the ring topology is equal for all initial individuals. However, in this case the links that belong to this ring are maintained at the crossover operation being changed only in the mutation operation, which penalize the obtained results.

B. Impact of the Combinations

In this section, the eight combinations are compared and analyzed. The presented results are obtained using an initial population generated using [24] and a uniform demand model. Figure 5 presents the best solution obtained in each iteration by the genetic algorithm, for each combination. Results for the vBNS network, a network with 12 nodes, are presented in Fig. 5(a). In Fig. 5(b) are presented the results for the SPAIN network, a network with 17 nodes. The lower bound obtained using the ILP model is also presented as a solid line.

As can be seen in Fig. 5(a), for the 12 node network, the convergence to a solution is fast for all combinations. The difference between the best solution obtained for each combination is not significant. As the number of nodes increases, the convergence to a solution is slower. As can be seen in Fig. 5(b), the convergence to a solution is only visible after the 60th iteration. The difference between the quality of the obtained solutions increases as well.

To compare the eight combinations the gap of the best solution obtained for each method, among the five runs, is presented in Table V. The best solution obtained is bold marked. The genetic algorithm with a population of 500 individuals, roulette wheel selection, and uniform crossover obtains the best solution for eight networks; see Table V. Moreover, a population of 500 individuals, tournament selection, and uniform crossover equals the best solution obtained in six networks. The second-best solution is also always obtained by one of these combinations. The difference between the solutions obtained by each combination ranges between 0% and 14%, tending to increase with the increase in the number of nodes.

Making a comparison among the combinations, the uniform crossover obtains better solutions than the single-point crossover, independently of the number of individuals and selection method; see Table V. The uniform crossover does not preserve large blocks of the progenitors' genetic code to the offspring; therefore it increases the population diversity
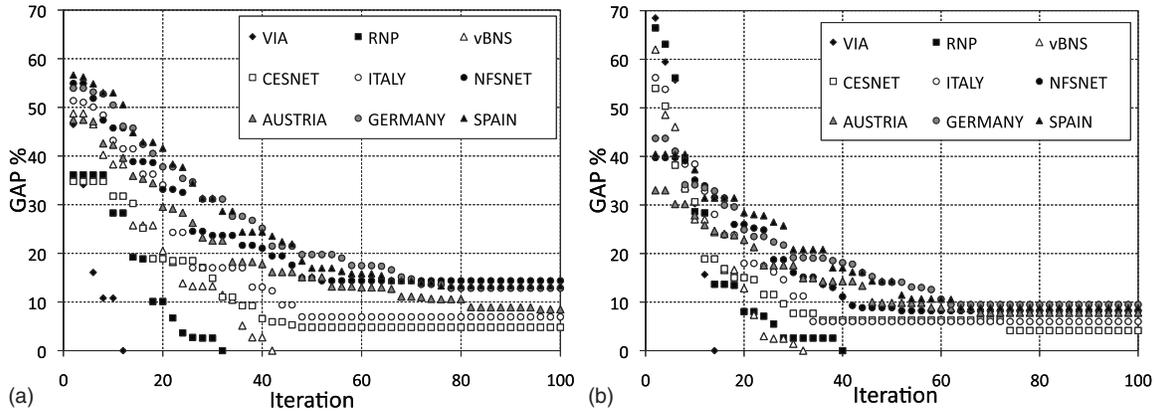
Fig. 4.   Evolution of the gap for the best solution obtained for the nine reference networks in each iteration for initial populations generated with the (a) random topology generator [17] and (b) topology generator presented in [24].

and allows the genetic algorithm to obtain better solutions. On the other hand, fixing the crossover operator, the solutions obtained using the roulette wheel selection method are quite similar to the ones obtained using the tournament method. Improvements are observed when using the roulette wheel selection method. Comparing the results obtained with the different sizes of population, fixing the selection method and crossover operator, large populations (500 individuals) obtain better solutions than smaller ones (100 individuals). However, runs were also done with populations of 1000 individuals and only residual improvements were obtained, with relation to solutions obtained with 500
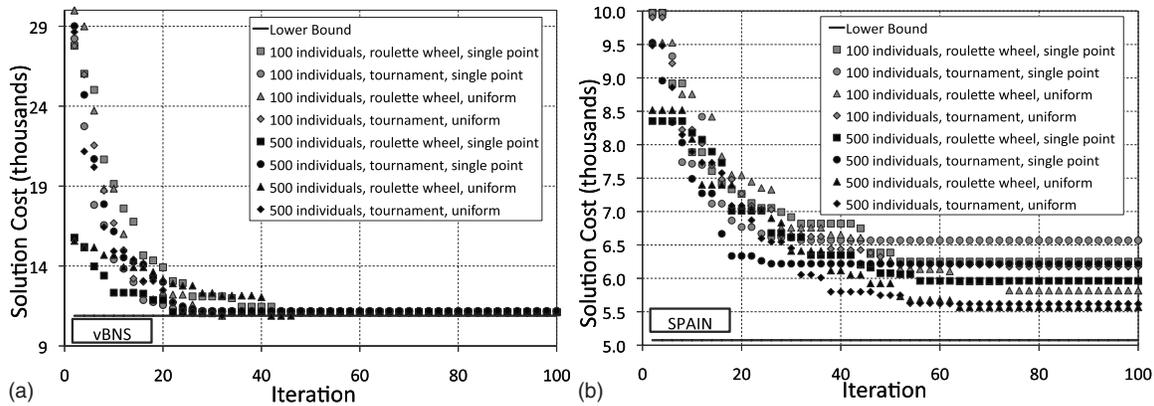

Fig. 5.   Evolution of the best solution obtained in each iteration using the topology generator presented in [24] for the eight considered combinations and lower bound obtained using the ILP model (solid black line) for (a) the vBNS network (12 nodes) and (b) the SPAIN network (17 nodes).

TABLE V

GAP OF THE BEST SOLUTION OBTAINED WITH EACH OF THE EIGHT CONSIDERED COMBINATIONS FOR INITIAL POPULATIONS GENERATED USING [24]

| | 100 Individuals | | | | 500 Individuals | | | |
| | Single Point | | Uniform | | Single Point | | Uniform | |
Network	Roulette Wheel	Tournament	Roulette Wheel	Tournament	Roulette Wheel	Tournament	Roulette Wheel	Tournament
VIA	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
RNP	2.6	7.9	3.5	2.7	1.9	1.9	**0.0**	**0.0**
vBNS	1.9	2.5	0.0	2.7	2.1	2.7	**0.0**	**0.0**
CESNET	6.0	8.9	4.1	6.0	7.0	7.3	4.8	**4.1**
ITALY	15.6	16.3	8.5	10.1	21.1	10.4	**6.0**	**6.0**
NFSNET	13.2	22.2	9.7	11.3	10.5	13.1	**8.2**	8.7
AUSTRIA	17.6	14.8	9.8	11.3	13.9	13.7	**7.8**	8.4
GERMANY	19.6	22.4	13.9	13.9	16.8	14.9	**9.5**	**9.5**
SPAIN	18.8	22.7	12.8	17.8	14.9	18.4	**8.8**	9.7

TABLE VI
COMPUTATIONAL RESULTS OBTAINED USING THE ILP MODEL AND THE GENETIC ALGORITHM FOR UNIFORM AND NONUNIFORM
DEMAND MATRICES

| Network | Nodes | Uniform Demand Model | | | | Nonuniform Demand Model | | | |
| | | ILP | | GA | | ILP | | GA | |
		Time	Gap	Time	Gap	Time	Gap	Time	Gap
VIA	9	1 s	0.0	8 s	0.0	4 s	0.0	8 s	0.0
RNP	10	42 s	0.0	27 s	0.0	24 m	0.0	27 s	3.2
vBNS	12	2 m	0.0	32 s	0.0	10 h	1.8	32 s	4.6
CESNET	12	7 h	0.0	01 m	4.1	10 h	0.3	01 m	5.7
ITALY	14	10 h	4.4	02 m	6.0	10 h	3.2	02 m	9.3
NFSNET	14	10 h	3.0	02 m	8.2	10 h	5.7	02 m	9.6
AUSTRIA	15	10 h	4.4	02 m	7.7	10 h	8.0	02 m	10.8
GERMANY	17	10 h	6.7	04 m	9.5	8 h (OM)	8.6	04 m	12.2
SPAIN	17	10 h	8.3	04 m	8.8	10 h	8.4	04 m	13.0

individuals.

In spite of the individual generation and the crossover operators being random, the difference between the solutions obtained by each run, among the five, is not significant. Moreover, as the size of the network increases, such difference decreases.

C. Impact of the Traffic Model

In the following, we analyze the best results obtained using the genetic algorithm and the results obtained using the ILP model, for uniform and nonuniform demand matrices. The nonuniform demand matrices are randomly generated with $0 \geqslant B_{od} \geqslant 5$. The gap and the processing time for the solutions obtained using the ILP model and the genetic algorithm (GA) are presented in Table VI.

Considering uniform demand matrices, the ILP model obtained the optimal solution for networks with fewer than 12 nodes. Considering the vBNS and CESNET networks, note that in spite of having the same number of nodes, the processing time required using the ILP model to achieve the optimum solution is substantially different; see Table VI. A reason for this may be found in the difference of the geographical area where the networks are implemented. The vBNS network, with 12 nodes, is in the USA and the CESNET network, also with 12 nodes, is in the Czech Republic. Due to the large area that the USA network has to cover, the majority of its links are fixed since the beginning due to the distance. For networks with more than 12 nodes the ILP model obtains a solution with a gap smaller than 8.3% within 10 hours. With the exception of the GERMANY and SPAIN networks, the ILP model obtains a solution with a gap smaller than 10% in less than 3 hours.

The genetic algorithm is much faster than the ILP model and obtains near-optimal solutions. For networks with less than 12 nodes the genetic algorithm obtains either the optimal solution or a solution with a gap of 4% within 1 minute. For networks with more than 12 nodes the genetic algorithm obtains solutions with gaps smaller than 10% within 4 minutes. A solution was obtained in approximately 45 minutes for a network with 100 nodes. In this case, the

gap was not calculated as this problem cannot be addressed using the ILP model within a reasonable time and computational effort.
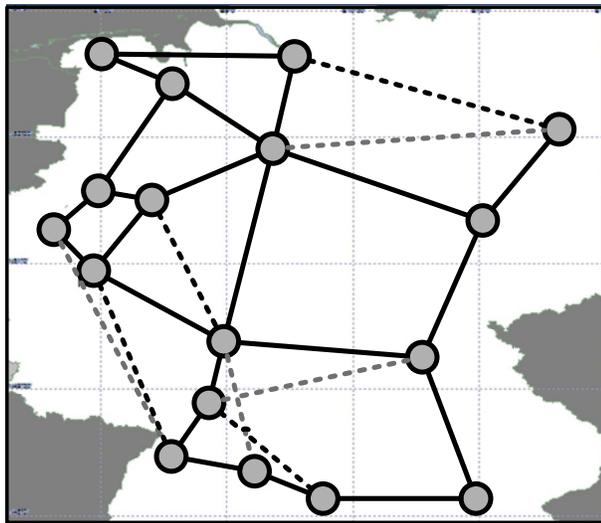
The results obtained using the ILP model and the genetic algorithm for the GERMANY network and a uniform demand matrix can be observed in Fig. 6(a). The dashed links are the ones that differ in both solutions. The black dashed lines represent the links of the topology obtained using the genetic algorithm. The gray dashed lines represent the links of the topology obtained using the ILP model. The black solid lines represent the common links to both solutions. None of the topologies obtained are optimal; see Table VI. However, the majority of the optimal links are already present in both solutions.

Considering nonuniform demand matrices, the complexity of the problem increases. In this case the ILP model only obtained the optimal solution, within the time limit, for networks with less than 10 nodes. However, solutions with a gap smaller than 8.6% can still be obtained in 10 hours. We also observe that the ILP model obtains solutions with gaps smaller than 10% within 3 hours of processing time. For the GERMANY network, the computer runs out of memory before the end of the 10 hours of processing time.

The genetic algorithm maintains the processing time, although the results obtained suffered an increase in the gap. We observed that such an increase is due to the routing algorithm. The optimal routing is not always the shortest path; sometimes longer routes can optimize the network available resources. Figure 6(b) depicts the best result obtained with the node location of the RNP network, using the ILP model and the genetic algorithm with a nonuniform demand matrix. As can be seen in Table VI, the genetic algorithm obtained a solution with a gap of 3.2%. Nevertheless, the physical topology obtained is the optimal; see Fig. 6(b). The gap is only due to suboptimal routing.

VI. CONCLUSIONS

We have developed a genetic algorithm for the design of minimum CAPEX survivable optical networks. The ratio be-

(a)
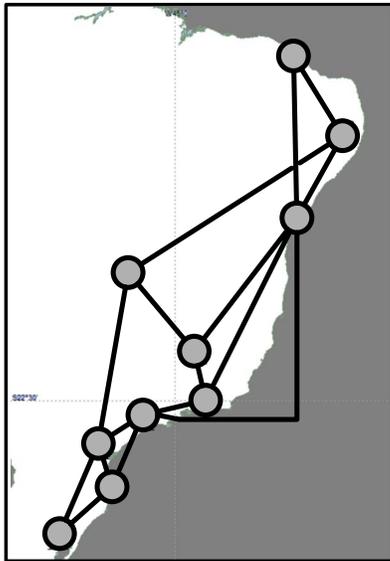
(b)

Fig. 6. (Color online) Topologies obtained using the ILP model and the genetic algorithm for the node location of (a) the GERMANY network with the uniform demand matrix and (b) the RNP network with the nonuniform demand matrix. The dashed links differ in both solutions and the solid links are common.

tween the gap (of the order of 10%) and the processing time obtained, using the genetic algorithm, encourage the use of this kind of heuristic within the survivable optical network design problem. An ILP model was also presented to evaluate the quality of the genetic algorithm solutions. Two initial population generators, two selection methods, two crossover operators, and two sizes of populations were compared within the genetic algorithm. Computational results obtained using the node location of real telecommunication networks showed that initial populations generated using a method that preserves the main characteristics of real optical networks improves the quality of the obtained solutions. Moreover, crossover operators that do not preserve large blocks of the genetic code increase the diversity of the population and the probability of finding better solutions.

REFERENCES

[1] S. K. Korotky, "Network global expectation model: a statistical formalism for quickly quantifying network needs and costs," *J. Lightwave Technol.*, vol. 22, no. 3, pp. 703–722, 2004.

[2] H. Kerivin and A. R. Mahjoub, "Design of survivable networks: a survey," *Networks*, vol. 46, no. 1, pp. 1–21, 2005.

[3] O. Klopfenstein, "Access network dimensioning with uncertain traffic forecasts," in *Proc. 13th Int. Telecommunications Network Strategy and Planning Symp.*, 2009, pp. 1–52.

[4] D. Jungnickel, *Graphs, Networks and Algorithms*. Springer, 2008.

[5] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

[6] S. Soni, R. Gupta, and H. Pirkul, "Survivable network design: the state of the art," *Inf. Syst. Front.*, vol. 1, no. 3, pp. 303–315, 1999.

[7] B. Caenegem, W. Parys, F. De Tuck, and P. Demeester, "Dimensioning of survivable WDM networks," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 7, pp. 1146–1157, 1998.

[8] S. Soni and H. Pirkul, "Design of survivable networks with connectivity requirements," *Telecommun. Syst.*, vol. 20, no. 1, pp. 133–149, 2002.

[9] A. Balakrishnan, T. L. Magnanti, and P. Mirchandani, "Connectivity-splitting models for survivable network design," *Networks*, vol. 43, no. 1, pp. 10–27, 2004.

[10] A. Balakrishnan, P. Mirchandani, and H. P. Natarajan, "Connectivity upgrade models for survivable network design," *Oper. Res.*, vol. 57, no. 1, pp. 170–186, 2009.

[11] C. Pluntke, M. Menth, and M. Duelli, "CAPEX-aware design of survivable DWDM mesh networks," in *Proc. IEEE Int. Conf. on Communications (ICC)*, 2009.

[12] M. Duelli, C. Pluntke, and M. Menth, "Minimizing installation costs of survivable DWDM-mesh networks: a heuristic approach," in *Proc. Next Generation Internet Networks (NGI 2008)*, 2008, pp. 15–22.

[13] A. Jarray, B. Jaumard, and A. C. Houle, "Minimum CAPEX/OPEX design of optical backbone networks," in *Proc. Int. Conf. on Ultra Modern Telecommunications & Workshops (ICUMT '09)*, 2009, pp. 1–8.

[14] I. de Miguel, R. Vallejos, A. Beghelli, and R. J. Durán, "Genetic algorithm for joint routing and dimensioning of dynamic WDM networks," *J. Opt. Commun. Netw.*, vol. 1, no. 7, pp. 608–621, 2010.

[15] L. S. Buriol, M. G. C. Resende, and M. Thorup, "Survivable IP network design with OSPF routing," *Networks*, vol. 49, pp. 51–64, 2007.

[16] D. A. R. Chaves, C. J. A. Bastos-Filho, and J. F. Martins-Filho, "Multiobjective physical topology design of all-optical networks considering QoS and Capex," in *Optical Fiber Communication Conf.*, 2010, paper JThA45.

[17] J. F. Labourdette, E. Bouillet, R. Ramamurthy, and A. A. Akyama, "Fast approximate dimensioning and performance analysis of mesh optical networks," *IEEE/ACM Trans. Netw.*, vol. 3, no. 4, pp. 906–917, 2005.

[18] E. Bouillet, G. Ellinas, J. F. Labourdette, and R. Ramamurthy, *Path Routing in Mesh Optical Networks*. Wiley, 2007.

[19] S. Ramamurthy, L. Sahasrabuddhe, and B. Mukherjee, "Sur-

vivable WDM mesh networks," *J. Lightwave Technol.*, vol. 21, no. 4, pp. 870–883, 2003.

[20] S. Azodomilky, M. Klinkowski, E. Marin, D. Careglio, J. S. Pareta, and I. Tomkos, "A survey on physical layer impairments aware routing and wavelength assignment algorithms in optical networks," *Comput. Netw.*, vol. 53, no. 7, pp. 926–944, 2008.

[21] R. Batchellor and O. Gerstel, "Cost effective architectures for core transport networks," in *Optical Fiber Communication Conf. and Expo. and the Nat. Fiber Optic Engineers Conf.*, 2006, paper PDP42.

[22] R. Ramaswami and K. N. Sivarajan, *Optical Networks: A Practical Perspective*. Morgan Kaufmann, 2002.

[23] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, 1989.

[24] C. Pavan, R. M. Morais, F. Rocha, and A. N. Pinto, "Generating realistic optical transport network topologies," *J. Opt. Commun. Netw.*, vol. 2, no. 1, pp. 80–90, 2010.

[25] B. Waxman, "Routing of multipoint connections," *IEEE J. Sel. Areas Commun.*, vol. 6, no. 9, pp. 1617–1622, Dec. 1988.

[26] M. Mitchell, *An Introduction to Genetic Algorithms*. MIT Press, 1996.

[27] R. Huelsermann, M. Gunkel, C. Meusberger, and D. A. Schupke, "Cost modeling and evaluation of capital expenditures in optical multilayer networks," *J. Opt. Netw.*, vol. 7, no. 9, pp. 814–833, 2008.

Rui Manuel Morais was born in Gouveia, Portugal. He graduated in applied mathematics and computation and obtained the M.Sc. degree in mathematics and applications, both from the University of Aveiro, Aveiro, Portugal, in 2006 and 2008, respectively. He joined Instituto de Telecomunicações in 2007 working on the dimensioning of optical networks. He is currently working toward his Ph.D. degree at the Department of Electronics, Telecommunications and Informatics, University of Aveiro, Portugal. His main research interests are the dimensioning problem with incomplete information and the optical network design problem.

Claunir Pavan received the B.Tech. degree in informatics technology from the West University of Santa Catarina—UNOESC, SC, Brazil, in 2000 and the M.Sc. degree in computer science from the Federal University of Santa Catarina—UFSC, SC, Brazil, in 2003. From 2003–2005 he was with the UNOESC, where he had been teaching courses on information systems and was head of the CTIC—Technology Information and Communications Department. From 2005, he moved to Portugal to work as a Researcher with the "Integration Strategies for IP over WDM Networks" project at the Institute of Telecommunications. He is currently working toward his Ph.D. degree at the Department of Electronic, Telecommunications and Informatics, University of Aveiro, Portugal. His main

research interest include the dimensioning of optical multilayer networks.

Armando Nolasco Pinto graduated in electronic and telecommunications engineering in 1994 and obtained the Ph.D. degree in electrical engineering in 1999, both from the University of Aveiro, Aveiro, Portugal. In 2000, he became an Assistant Professor in the Electrical, Telecommunications and Informatics Department, University of Aveiro, and a Researcher at the Institute of Telecommunications, Aveiro. During the academic year of 2006–2007 he was a Visiting Professor at the Institute of Optics, University of Rochester, Rochester, NY. At present, he leads a research group at the Institute of Telecommunications focusing on high-speed optical communication systems and networks. He has published more than 100 scientific papers in international journals and conferences. Dr. Pinto is a member of the Optical Society of America (OSA) and a senior member of the Institute of Electrical and Electronics Engineers (IEEE).

Cristina Requejo is an Assistant Professor in the Department of Mathematics, University of Aveiro, Portugal. She holds a Ph.D. in operations research from the University of Lisbon, Portugal. Her main research interests are discrete and combinatorial optimization, including integer programming problems with a particular emphasis on network design problems. Telecommunication networks have been central to the applied part of her research work.