

Topology Design of Transparent Optical Networks Resilient to Multiple Node Failures

Fábio Barbosa
Instituto de Telecomunicações
Universidade de Aveiro
3810-193 Aveiro, Portugal
fabio Barbosa@ua.pt

Amaro de Sousa
Instituto de Telecomunicações
Universidade de Aveiro
3810-193 Aveiro, Portugal
asou@ua.pt

Agostinho Agra
CIDMA, Dept. Matemática
Universidade de Aveiro
3810-193 Aveiro, Portugal
aagra@ua.pt

Abstract—Consider the resilience of a network defined by the average 2-terminal reliability (A2TR) against a set of critical node failures. Consider an existing transparent optical network with a total fibre length L . The first goal of this paper is to assess the resiliency gap between the existing topology and a new network topology designed to maximize its resilience with the same fibre budget L . The resiliency gap gives us a measure of how good the resilience of existing network topologies are. Consider now that an existing network is upgraded with new links aiming to maximize its resiliency improvement with a fibre budget L' . The second goal of this paper is to assess how much the resiliency gap can be reduced between a good upgraded solution and a network topology designed to maximize its resiliency with the same fibre budget $L + L'$. The gap reduction gives us a measure of how close to the best resilience the upgraded solutions can get for different values of L' .

To reach these goals, we first describe how the Critical Node Detection problem is defined and solved in the context of transparent optical networks. Then, we propose a multi-start greedy randomized method to generate network topologies, with a given fibre length budget, that are resilient to critical node failures. This method is also adapted to the upgrade of an existing network topology. At the end, we run the proposed methods on network topologies with public available information. The computational results show that the resiliency gap of existing topologies is significantly large but network upgrades with $L' = 10\%L$ can significantly reduce the resiliency gaps provided that such upgrades are aimed at maximizing the network resilience to multiple node failures.

Index Terms—Transparent Optical Networks; Critical Node Detection; Resilient Network Design; Disasters

I. INTRODUCTION

Large-scale failures can seriously disrupt any telecommunication network due to either natural, technological or malicious human activities [1] (two surveys conducted within COST Action RECODIS are [2] on strategies to protect networks against large-scale natural disasters and [3] on security challenges in communication networks). So, an emerging research topic is the design of telecommunication networks enhancing their resilience to large-scale failures. To reach this goal, we must first adopt a proper resiliency evaluation metric and, then, we must investigate proper network design methods aiming to maximize the network resiliency metric to large-scale failures.

This work addresses the design of resilient network topologies in the context of transparent optical networks. Note that,

in general, multiple failures might involve only links or nodes and links (a node failure implies that its links also fail). For example, in malicious human attacks, node shutdowns are harder to realize but are the most rewarding in the attackers perspective (the shutdown of a single node is also able to shut down multiple links). Node failures are more harmful to the resilience of networks and, so, we address the topology design of transparent optical networks which must be resilient to multiple node failures.

For a given topology, if some nodes are considered critical due to some reason, the network design should take this into consideration, as in [4] where the approach proposed in [5] is adapted to the design of a transparent optical network minimizing the failure impact of a given set of critical nodes. Here, we consider the resiliency metric defined by the average 2-terminal reliability (A2TR) and, for a given network topology, we evaluate this metric against a set of critical node failures. A2TR is defined as the number of node pairs that remain connected if all critical nodes fail and the set of critical nodes is the optimal solution of a Critical Node Detection (CND) optimization problem.

CND problems have been considered in different contexts and are gaining special attention in the vulnerability evaluation of telecommunication networks to large-scale failures [2]. In [6], CND is defined as the detection of a given number c of critical nodes aiming to minimize the number of connected node pairs. More recently, this and other variants of CND have also been addressed [7]–[10] but none of these works addresses the CND problem in the context of transparent optical networks.

In these networks, data is converted into htwaveht in the source node and transmitted through an all optical path, named *lightpath*, towards the destination node. Due to many optical degradation factors, like attenuation, dispersion, crosstalk and other non-linear factors, there is a maximum length, named *transparent reach*, for each lightpath to work properly. Moreover, the length of a path depends both on the length of its links and on its number of hops. The optical degradation suffered by a lightpath while traversing an intermediate node is usually modelled by a given fibre length value d , *i.e.*, by considering it equivalent to the degradation incurred due to the transmission over a given fibre of length d . So, when accounting the A2TR

metric, the CND problem has to consider that two nodes are connected only if the surviving network provides it with a shortest path within the transparent reach. Here, a proper Integer Linear Programming (ILP) description of this CND problem variant is provided together with a row generation approach to compute its optimal solution.

Other metrics have been used to evaluate the vulnerability of networks in other contexts [11] or assuming multiple failures with geographical correlation between failing elements [12]. There are also works on improving the preparedness of networks to multiple failures, some by changing the network topology [13], [14], [15], while others by proposing strategies to recover from failures [16], [17]. None of these works, though, uses the optimal solution of CND to assess the vulnerability of networks. On the other hand, in [18], CND is used but resiliency improvement is exploited by optimal robust node selection on a given topology. The advantage of using CND is that it provides a worst case resiliency analysis, *i.e.*, in any failure involving the same number of failing nodes, the resulting A2TR is never worse than the value provided by the solution of CND.

Here, we propose a multi-start greedy randomized method to generate network topologies, with a given fibre length budget, that are resilient to critical node failures. The method is also adapted to the upgrade of an existing topology. For an existing network with a total fibre length L , the first aim is to assess the resiliency gap between the existing topology and a new network topology designed to maximize its resilience with the same fibre budget L . If the existing network is to be upgraded with new links within a fibre budget L' , the second aim is to assess how much the resiliency gap can be reduced between a good upgraded topology and a network topology designed to maximize its resiliency with the same fibre budget $L + L'$.

The paper is organized as follows. Section II describes a path-based Mixed ILP (MILP) model defining the CND problem, a row generation approach used to solve it and centrality based heuristics combined with a local search method to approximate it. Section III proposes the multi-start greedy randomized method to generate network topologies resilient to critical node failures. The computational results are presented and discussed in Section IV. Finally, Section V presents the main conclusions of this work.

II. CRITICAL NODE DETECTION (CND) PROBLEM

Consider a transparent optical network represented by an undirected graph $G = (N, E)$ where $N = \{1, \dots, n\}$ is the set of nodes and $E \subseteq \{(i, j) \in N \times N : i < j\}$ is the set of fibre links. For each link $(i, j) \in E$, parameter l_{ij} represents its length.

The transparent reach of the network is denoted by parameter $T > 0$ and the fibre length equivalent to the degradation suffered by a lightpath while traversing an intermediate node is denoted by parameter $d > 0$. We assume that $l_{ij} \leq T$ for all $(i, j) \in E$; otherwise, such link is worthless and can be removed from G .

The set of all paths in G between $i \in N$ and $j \in N$ (with $i < j$ and $(i, j) \notin E$) with length not greater than T is denoted by P_{ij} . Each path $p \in P_{ij}$ is defined by the binary parameters β_k^p , indicating whether node k (which can be an end node) is in p or not, and α_{kt}^p indicating whether link (k, t) , $k < t$ is in p or not. So, P_{ij} is composed by all paths p such that

$$\sum_{k=1}^{n-1} \sum_{t=k+1}^n \alpha_{kt}^p l_{kt} + d \left(\sum_{k=1}^n \beta_k^p - 2 \right) \leq T.$$

A. Path-based MILP model

For each node $i \in N$, we consider a binary variable v_i indicating whether i is a critical node or not. For each node pair (i, j) , with $i, j \in N : i < j$, the binary variable u_{ij} is 1 if nodes i and j are connected through a path satisfying the transparent reach T , and 0 otherwise.

Then, for a given number $c \in \mathbb{N}$ of critical nodes, a path formulation for the CND problem is given by the following ILP model.

$$\min \quad z := \sum_{i=1}^{n-1} \sum_{j=i+1}^n u_{ij} \quad (1)$$

$$s.t. \quad \sum_{i=1}^n v_i \leq c, \quad (2)$$

$$u_{ij} + v_i + v_j \geq 1, \quad (i, j) \in E, \quad (3)$$

$$u_{ij} + \sum_{k=1}^n \beta_k^p v_k \geq 1, \quad (i, j) \notin E, p \in P_{ij}, \quad (4)$$

$$v_i \in \{0, 1\}, \quad i \in N, \quad (5)$$

$$u_{ij} \in \{0, 1\}, \quad i, j \in N : i < j. \quad (6)$$

The objective (1) is to minimize z defined as the total number of connected node pairs in the surviving graph (*i.e.* the graph given by removing all critical nodes from G). Constraint (2) ensures that at most c nodes are selected as critical nodes (in any optimal solution, c nodes are selected). Constraints (3) guarantee that a pair of adjacent nodes is connected if none of the two nodes is a critical node. Constraints (4) are the generalization of constraints (3) for the node pairs that are not adjacent in G : node pair (i, j) is connected if there is one path $p \in P_{ij}$ such that none of its nodes is a critical node. Constraints (5)-(6) are the variable domain constraints.

Note that, since variables v_i are binary, constrains (3)–(4) impose $u_{ij} \geq 1$ when nodes i and j are connected, which then, due to the objective function, forces $u_{ij} = 1$. Therefore, constraints (6) can be replaced by $u_{ij} \geq 0$. The resulting Mixed Integer Linear Programming (MILP) model will be considered henceforward.

B. Row generation approach

The exact number of constraints (4) of the MILP model depends on the graph topology, the link lengths and the values of T and d . However, the model becomes too large for relative small sized instances. Here, we propose a row generation approach to solve it. The exact algorithm is described in Algorithm 1.

Initially, inequalities (4) are ignored and the relaxed MILP problem is solved. Then, the separation problem associated with inequalities (4) is solved, all violated inequalities are added to the model and the MILP is solved again. The process is repeated until no violated inequality is found.

The separation problem associated with constraints (4) is solved in the following way. We determine a subgraph G^C removing from G the critical nodes and the corresponding incident edges ($G^C = (N \setminus C, E^C)$) and adding d to the length of each edge in E^C . Note that the number of intermediate nodes of a path is equal to the number of edges minus one. As a consequence, the shortest path value in G^C is equal to the path length plus d . So, we determine the shortest path in G^C between all pairs of nodes i and j in $N \setminus C$, such that $(i, j) \notin E^C$, using Dijkstra algorithm, and each shortest path whose length is not higher than $T + d$ is used to generate a new inequality (4) that is added to the model.

Algorithm 1 Exact algorithm for the CND problem

- 1: Solve the MILP model without constraints (4); let (u^*, v^*) be the optimal solution
 - 2: **repeat**
 - 3: Set NCuts $\leftarrow 0$ and $C \leftarrow \{i \in N : v_i^* = 1\}$
 - 4: Compute subgraph $G^C = (N \setminus C, E^C)$ where $E^C = \{(i, j) \in E : i, j \notin C\}$
 - 5: **for all** node pair $(i, j) \notin E^C$ with $i < j$ **do**
 - 6: Run Dijkstra algorithm (adding d to the length of each edge) to find the shortest path $p_{ij} \in P_{ij}$ and its length d_{ij}
 - 7: **if** $d_{ij} \leq T + d$ and $u_{ij}^* + \sum_{k=1}^n \beta_k^{p_{ij}} v_k^* = 0$ **then**
 - 8: Add constraint (4) corresponding to path p_{ij}
 - 9: NCuts \leftarrow NCuts + 1
 - 10: **end if**
 - 11: **end for**
 - 12: **if** NCuts > 0 **then**
 - 13: Solve MILP model with the added constraints. Update (u^*, v^*)
 - 14: **end if**
 - 15: **until** NCuts = 0
-

C. Centrality based heuristics

Heuristic methods based on centrality measures can be used to compute critical node sets because they run very quickly although not providing optimal solutions.

Algorithm 2 presents a general heuristic framework for using these measures: in each iteration of the **For** cycle, a node is selected according to the centrality measure chosen (steps 3 and 4) and removed from the graph (step 5). These heuristics will be used later on in the network design task as a means to shorten the evaluation runtime of solutions. Preliminary tests have shown that these heuristics are worthwhile with the following centrality measures:

- Degree centrality. The central node in step 3 is the node with highest degree in the current graph G' .

- Betweenness centrality. In the current graph G' , the betweenness of node i is the number of shortest paths (adding d for each intermediate node) between all nodes with length not greater than T that include node i as an intermediate node. The central node in step 3 is the node with highest betweenness.

Algorithm 2 Iterative heuristic based on a centrality measure

- 1: Set $C \leftarrow \emptyset$ and $G' \leftarrow (N, E)$
 - 2: **for all** $k = 1$ to c **do**
 - 3: Select the central node $i \in N$ of graph G'
 - 4: Set $C \leftarrow C \cup \{i\}$
 - 5: Remove from G' node i and all its incident edges
 - 6: **end for**
-

D. Local search approach

Note that, for a given set of nodes $C \subset N$ with $|C| = c$, we can compute in polynomial time its CND value z by determining the total number of shortest paths with length not higher than T between all node pairs in the surviving graph (*i.e.*, the graph that results from G by removing the set of nodes C and corresponding incident edges). So, in order to potentially improve the solutions obtained with the previous heuristics, we also consider a node based local search method (described in Algorithm 3) that evaluates each swap of a critical node by a non-critical neighbour node in G .

Algorithm 3 Local Search Method

- 1: Given a critical node set $C \subset N$ with $|C| = c$ and its CND value z
 - 2: **repeat**
 - 3: **for all** $i \in C, j \in N \setminus C : (\min(i, j), \max(i, j)) \in E$ **do**
 - 4: $C_i^j \leftarrow (C \setminus \{i\}) \cup \{j\}$ and compute its CND value z_i^j
 - 5: **end for**
 - 6: **if** $\min\{z_i^j\} < z$ **then**
 - 7: Update $z \leftarrow \min\{z_i^j\}$, and $C \leftarrow C_i^j$ accordingly
 - 8: **end if**
 - 9: **until** z is not updated
-

III. NETWORK DESIGN PROBLEM

In this section, we propose a multi-start greedy randomized algorithm to generate network topologies, with a fibre length budget given by B , that are resilient to critical node failures. In the proposed algorithm, the evaluation of each network topology uses the methods described in the previous section.

In general, a greedy randomized algorithm builds a network topology by starting with a graph with an empty set of fibre links $G = (N, \emptyset)$ and randomly selecting one link at a time until no new link can be added within the given budget B .

A key issue of this approach is how to define the probability $P((i, j))$ of each new link (i, j) , with $i < j$, being selected so that the method can efficiently find good network topologies. After testing multiple strategies, the best results were obtained

by guaranteeing that at least one end node of each new link is one of the lowest degree nodes of the current partial topology and by giving an higher probability to shorter links.

After fine-tuning, the best algorithm was obtained considering the probabilities as follows. First, consider that at each step the set of already selected links is E , δ_i is the degree of node i in $G = (N, E)$ and the remaining budget is $B_R = B - \sum_{(i,j) \in E} l_{ij}$. Then, for all node pairs $(i, j) \notin E$ such that $l_{ij} \leq B_R$ and at least one of the nodes (i or j) has the lowest degree in $G = (N, E)$ (i.e., $\min\{\delta_i, \delta_j\} = \min\{\delta_k : k \in N\}$), the probability is:

$$P((i, j)) = \frac{1}{(|\delta_i - \delta_j| + 1) l_{ij}^2} \quad (7)$$

while for all other node pairs (i, j) , $P((i, j)) = 0$.

Nevertheless, starting from an empty set of fibre links still did not allow to reach an efficient algorithm. Instead, we have investigated different criteria to adopt an initial non-empty set E_0 of fibre links. The most efficient algorithm was obtained by using the Relative Neighbourhood Graph (RNG) [19] as E_0 which is defined as follows: nodes $i, j \in N$ are connected by a link if and only if there is no other node $k \in N \setminus \{i, j\}$ such that $l_{ik} \leq l_{ij}$ and $l_{jk} \leq l_{ij}$. Our preliminary tests have shown that this graph provides a good initial balance between connectivity and amount of used fibre.

The resulting algorithm is described in Algorithm 4. Note that this algorithm can be easily adapted to the upgrade of an existing network topology by setting E_0 in step 1 with the link set of the existing topology instead of using the RNG.

Algorithm 4 Greedy Randomized Generation

- 1: Compute initial graph $G = (N, E_0)$
 - 2: Set $B_R \leftarrow B - \sum_{(i,j) \in E} l_{ij}$
 - 3: **repeat**
 - 4: Select a new link (i, j) with probabilities given by (7)
 - 5: $E \leftarrow E \cup \{(i, j)\}$
 - 6: $B_R \leftarrow B_R - l_{ij}$
 - 7: **until** $P((i, j)) = 0$, for all $i, j \in N : i < j$
-

Multiple runs of Algorithm 4 generate different topologies. So, in a multi-start greedy randomized algorithm, we run multiple times Algorithm 4, evaluate the CNL value z of each generated topology and store the topology with the highest z among all. The resulting algorithm is presented in Algorithm 5 with a stopping criteria given by maximum runtime.

Depending on the purpose of the algorithm, the initial topology $\bar{G} = (N, \bar{E})$ is set differently in step 1. When the algorithm is used to upgrade an existing topology, the initial topology is set to $\bar{G} = (N, \emptyset)$ with its CNL value $\bar{z} = 0$. When the algorithm is used to generate a topology better than a given one defined by a graph G and with a CNL value z , then, \bar{G} is set to G and its CNL value \bar{z} is set to z .

Recall that in the design of transparent optical networks, a topology is only valid if it is optically transparent, i.e., if the shortest path (adding d for each intermediate node) between

each node pair is not higher than T for all node pairs. So, each topology generated in step 3 is first validated in step 4 and discarded before evaluation if it is not optically transparent. Moreover, when the initial topology \bar{G} is 2-connected, we also require the solution of the algorithm to be 2-connected and discard the topologies accordingly. In the context of transparent optical networks, a topology is 2-connected if it is optically transparent for every removal of a single node.

In steps 5-19, each valid topology is evaluated saving as best topology the solution with the highest CNL value \bar{z} . Note that the most time consuming part of Algorithm 5 is the evaluation. The rationale of this algorithm is to use the heuristics described in the previous section to evaluate each generated topology and discard it whenever its objective value is lower than the current best solution \bar{z} . As a consequence, the exact method to detect the critical nodes only runs if none of the heuristics discard the topology under evaluation. Moreover, they are run from the fastest (Degree centrality), in terms of runtime, to the most time consuming (Exact CNL method).

Algorithm 5 Multi-Start Greedy Randomized Algorithm

- 1: Initialize $\bar{G} = (N, \bar{E})$ and its CNL value \bar{z}
 - 2: **repeat**
 - 3: Generate a new graph $G = (N, E)$ using Algorithm 4.
 - 4: **if** G is a valid topology **then**
 - 5: Run Algorithm 2, using Degree centrality
 - 6: Compute CNL value z_{Deg} from that node set
 - 7: **if** $z_{Deg} \geq \bar{z}$ **then**
 - 8: Run Algorithm 2, using Betweenness centrality
 - 9: Compute CNL value z_{Bet} from that node set
 - 10: **if** $z_{Bet} \geq \bar{z}$ **then**
 - 11: Run Algorithm 3, using node set corresponding to $\min\{z_{Deg}, z_{Bet}\}$, and compute z_{LS}
 - 12: **if** $z_{LS} \geq \bar{z}$ **then**
 - 13: Run Algorithm 1, obtaining z_{MLP}
 - 14: **if** $z_{MLP} \geq \bar{z}$ **then**
 - 15: $\bar{G} \leftarrow G$ and $\bar{z} \leftarrow z_{MLP}$
 - 16: **end if**
 - 17: **end if**
 - 18: **end if**
 - 19: **end if**
 - 20: **end if**
 - 21: **until** maximum runtime reached
-

IV. COMPUTATIONAL RESULTS

All computational results were obtained using the optimization software *Gurobi Optimizer* version 7.5.1, with programming language *Julia* version 0.6.0, running on a PC with an Intel Core i7, 2.3 GHz and 6 GB RAM. Following [22], we have assumed a transparent reach $T = 2000$ km corresponding to the use of OTU-4 lightpaths with a demand capacity of 100 Gbps. Moreover, we have considered $d = 60$ km.

The network topologies selected in our computational experiments are all optically transparent for $T = 2000$ km and are: Germany50 [20], PalmettoNet [21] and Missouri Network

Alliance (MissouriNA) [21]. Table I presents their topology characteristics in terms of number of nodes $|N|$ and fibre links $|E|$, total number of node pairs, minimum (δ_{\min}), average ($\bar{\delta}$) and maximum (δ_{\max}) node degree and an indication (in column '2-C') if the topology is (or is not) 2-connected.

Table I: Topology characteristics of each network.

Network	$ N $	$ E $	Pairs	δ_{\min}	$\bar{\delta}$	δ_{\max}	2-C
Germany50	50	88	1225	2	3.52	5	Yes
PalmettoNet	45	64	990	1	2.84	5	No
MissouriNA	64	80	2016	1	2.50	5	No

In all cases, the geographical location of nodes is publicly available but the geographical routes of fibre links is not known. So, we have considered that each link follows the shortest path over the surface of a sphere representing Earth. Table II presents the resulting length characteristics in terms of minimum (l_{\min}), average (\bar{l}), maximum (l_{\max}) and total (L) link length, and diameter, *i.e.*, the highest length among the shortest paths (adding d for each intermediate node) of all node pairs (all topologies are optically transparent for $T = 2000$ km since all diameter values are below 2000).

Table II: Length characteristics (in km) of each network.

Network	l_{\min}	\bar{l}	l_{\max}	L	Diameter
Germany50	26	100.7	252	8859	1417
PalmettoNet	19	67.0	177	4286	1298
MissouriNA	7	50.0	307	4001	1301

In the computational experiments, we have considered $c \in \{2, 3, 4, 5, 6\}$ as the number of critical nodes used to compute the resiliency metric z of each topology. For each network and each c , we started by computing (with Algorithm 5) a topology with a fibre budget B equal to the total fibre length L of the original topology. Then, we computed an upgraded topology for each original topology assuming a fibre budget $L' = p \times L$ with $p = 10\%$ and 20% . Finally, we computed a topology with a fibre budget $B = L + p \times L$ also for $p = 10\%$ and 20% . In each case, we gave a runtime limit of 5 hours to Algorithm 5.

Table III presents the resiliency value z of the best topologies obtained by the multi-start greedy randomized algorithm. Rows 'Original' refer to the original topologies (in column '0%') and upgraded topologies (in columns '10%' and '20%') while rows 'Generated' refer to the best topology solutions with a fibre budget $B = L + p \times L$ with $p = 0\%$, 10% and 20% . For each case, columns 'UB' presents the trivial upper bound of z given by the number of pairs of $|N| - c$ surviving nodes.

The first observation of these results is that the resiliency values are lower for higher number of critical nodes c , which is without surprise since more node failures disrupt an higher percentage of the network. Moreover, the resilience of the upgraded topologies is always significantly better for higher budget value L' . Finally, the best topologies are always

significantly better than the original/upgraded ones for PalmettoNet and MissouriNA. Nevertheless, this is not the case for Germany50 where the difference between the two types of solutions is already small for higher values of c and even null for many cases of the lower values of c . So, one major conclusion is that Germany50 is significantly more resilient to critical node failures than PalmettoNet and MissouriNA. To understand this fact, recall from the topology characteristics of the different networks (Table I) that Germany50 is the topology with the highest average node degree and the only one which is 2-connected. These two characteristics make this network more resilient than the two other networks.

More important than analysing the absolute resiliency values z , we need to analyse the resiliency gap between the original/upgraded topologies and the best topologies computed with the same fibre budget values. Figure 1 plots in a bar chart these gaps, for all networks and all values of c , computed as $\frac{z_B - z_{OU}}{z_B}$ where z_B is the resiliency value of the best topology and z_{OU} is the resiliency value of the original/upgraded topology. Blue bars present the resiliency gap between the best topology and the original topology. The resiliency gaps between the best topologies and the upgraded topologies are presented in the purple and green bars for $p = 10\%$ and 20% , respectively.

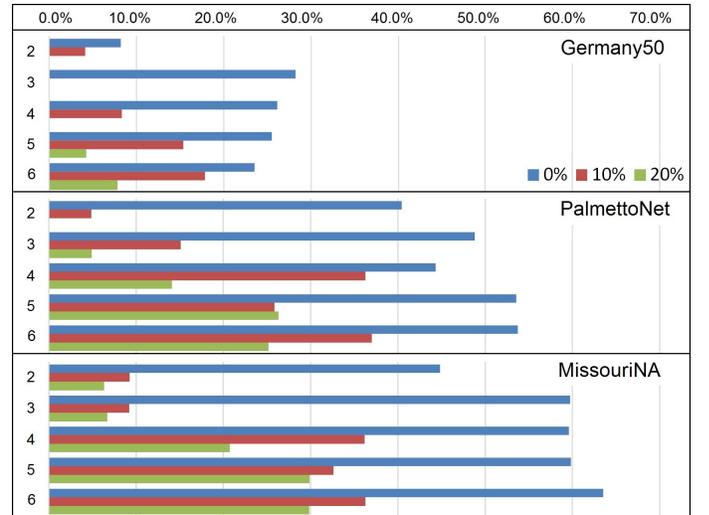


Figure 1: Resiliency gaps $\frac{z_B - z_{OU}}{z_B}$ (%) of all cases.

The blue bars of Figure 1 show that the resiliency gaps are lower for Germany50 (but still significant for a number of critical nodes $c \geq 3$) and very large for PalmettoNet and MissouriNA. These results reinforce the previous conclusion that Germany50 is more resilient than the others but also show that, in all cases, existing network topologies are not resilient to critical node failures. On the other hand, the resiliency gaps shown in the purple bars (corresponding to topology designs with 10% more total fibre length) represent, in all cases, a significant gap reduction when compared with the blue bars. This means that in all topologies and for all considered number of critical nodes, adding new links to an existing topology with a fibre budget of 10% enables solutions whose resiliency to

Table III: Resiliency value z of all cases obtained by the multi-start greedy randomized algorithm.

c	Network Instance	Germany50				PalmettoNet				MissouriNA			
		0%	10%	20%	UB	0%	10%	20%	UB	0%	10%	20%	UB
2	Original	1036	1081	1128		513	821	861		946	1555	1659	
	Generated	1128	1128	1128	1128	861	861	861	903	1714	1714	1771	1891
3	Original	711	991	1035		346	616	709		602	1362	1446	
	Generated	991	991	1035	1081	676	709	744	861	1495	1500	1550	1830
4	Original	640	830	906		284	427	510		455	762	1039	
	Generated	867	906	906	1035	510	582	582	820	1126	1194	1311	1770
5	Original	496	640	756		176	325	380		338	618	758	
	Generated	666	756	790	990	379	409	480	780	841	917	1081	1711
6	Original	415	498	606		123	235	286		253	457	550	
	Generated	543	606	658	946	266	322	358	741	694	717	784	1653

critical node failures becomes closer to a topology designed to maximize this resilience. Interestingly, the results of the green bars (corresponding to topology designs with 20% more total fibre length) are mixed, *i.e.*, in some cases, the additional 10% fibre budget enables a significant gap reduction while in other cases, the reduction is negligible.

Finally, we can distinguish two groups of results. For a number of critical nodes $c \leq 3$, the additional fibre budget of 20% makes in all networks the resiliency gap to become very small. For a number of critical nodes $c \geq 4$, and in the less resilient PalmettoNet and MissouriNA networks, the additional fibre budget of 20% is still not enough to make the resiliency gap small. This means that more fibre links are required in the upgrade of existing networks to reach the best resiliency to higher number of critical nodes.

Table IV presents, for each tested instance, the percentage of the total fibre length L of the original topology that is common to the best topology computed with the same fibre budget L . These results show that these percentage values are around 50%, with some small differences, for all topologies and all values of c , showing that the best topologies, in terms of resiliency to multiple node failures, are significantly different from the existing ones.

Table IV: Percentage of the total fibre length of the original topology common to the best topology.

c	Germany50	PalmettoNet	MissouriNA
2	43,6%	52,6%	47,1%
3	49,1%	52,1%	50,2%
4	45,8%	52,7%	49,5%
5	43,5%	51,1%	47,1%
6	46,3%	51,3%	48,0%
Average	45,7%	52,0%	48,4%

For illustrative purposes, Figure 2 presents the original topologies and the best topologies with the same fibre budget L obtained for $c = 3$ critical nodes. To understand the differences, links of the best topology not in the original topology are highlighted in dashed blue and, in both cases, critical nodes are represented with red squares. Also, Figure 3 presents the best upgraded solutions with $L' = 10\%L$ and $20\%L$ obtained also for $c = 3$ with the additional links

highlighted in dashed blue (again, critical nodes represented with red squares). The analysis of these topologies show that:

Germany50: The critical node set splits the original network in two components (10 and 37 nodes each) while it only isolates two nodes from the others in the best topology. Moreover, the critical node set isolates 2 nodes from the others in the 10% upgraded topology and a single node in the 20% upgraded topology.

PalmettoNet: The critical node set splits the original network in three components (6, 13 and 23 nodes each) while it splits the best topology in only two components (5 and 37 nodes each). Moreover, the critical node set splits the 10% upgraded topology in two components (7 and 35 nodes) and the 20% upgraded topology in two components (4 and 38 nodes). In this case, both the best topology and the two upgraded topologies are 2-connected.

MissouriNA: The critical node set splits the original network in three components (17, 20 and 24 nodes each) while it splits the best topology in three components (1, 5 and 55 nodes each). Moreover, the critical node set splits the 10% upgraded topology in two components (9 and 52 nodes) and the 20% upgraded topology in two components (6 and 55 nodes). In this case, the 20% upgraded topology is 2-connected but neither the best topology nor the 10% upgraded topology are, showing that the original MissouriNA is much less connected and, therefore, requires more fibre length upgrades to become 2-connected.

This analysis clearly highlights that the best topologies with the same total fibre of existing ones are much more resilient to critical node failures and the resiliency of existing topologies can be improved with the addition of new links.

Another aspect of interest is the comparison of the node degree distributions between the original topologies and the best topologies with the same total fibre. Figure 4 shows these distributions for the three network cases with the best topologies obtained for $c = 3$ critical nodes (original topologies in black and best topologies in blue). Interestingly, in the best topologies, there is a decrease of the number of nodes with the lowest and highest degrees and an increase of the number of nodes with degrees closer to the average. This observation also

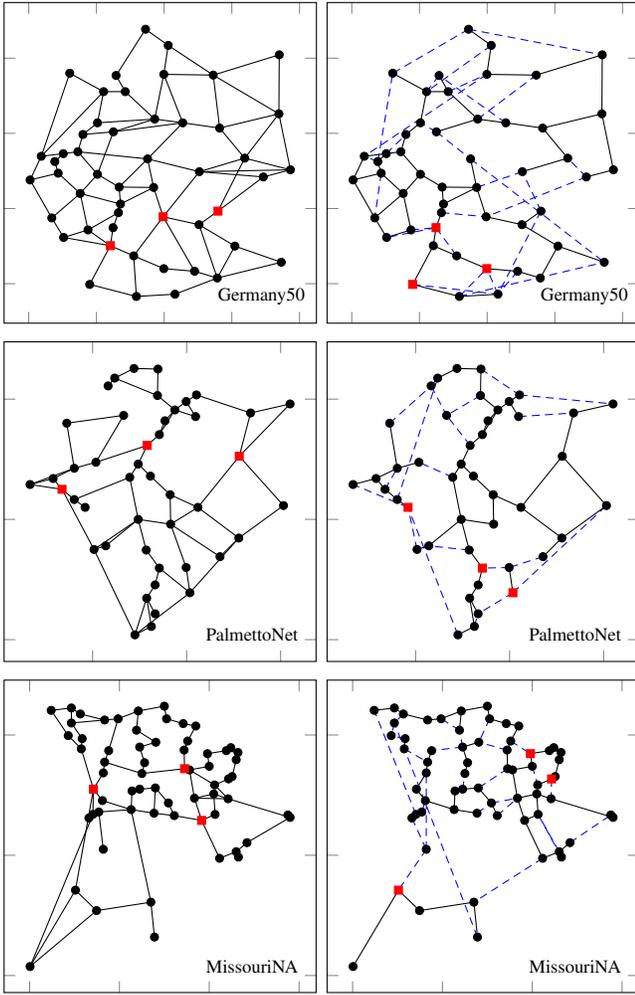


Figure 2: Original topologies (left) and best topologies (right) for $c = 3$. Links not in the original topology highlighted in dashed blue in the best topology (critical nodes in red squares).

stands in the best topologies for the other values of c showing that resilient topologies tend to have more homogeneous node degrees.

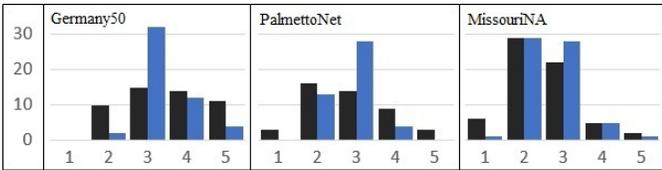


Figure 4: Node degree histograms of original topology (in black) and the best topology (in blue) for $c = 3$.

Finally, recall that Algorithm 5 (see Section III) uses heuristics in the evaluation of the CND value z of each valid topology as a means to minimize the number of times the exact method is used. In order to evaluate the efficiency of this strategy, Table V presents the average percentage of valid solutions that were discarded by the heuristics, row 'Success

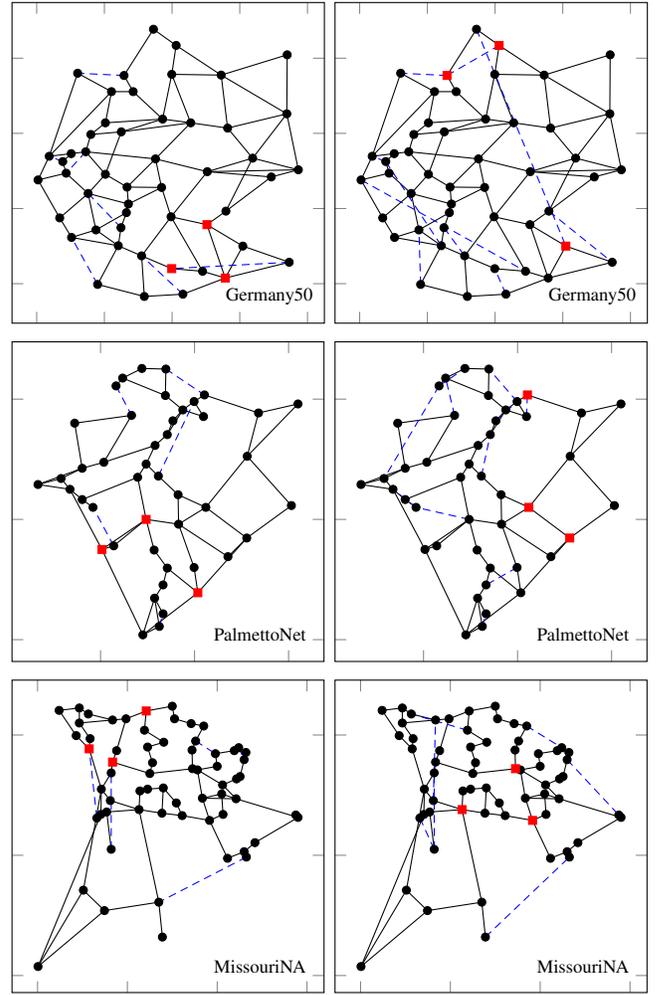


Figure 3: Best upgraded topologies with $L' = 10\%L$ (left) and $20\%L$ (right) for $c = 3$. Links added to the original topologies highlighted in dashed blue (critical nodes in red squares).

(%)', and the average runtime percentage the algorithm has spent while running the heuristics, row 'Success (%)', among all cases of each network topology and also among all cases of all topologies (column 'Average').

Table V: Average percentage of discarded solutions and average runtime percentage of the heuristics running Algorithm 5.

	Germany50	PalmettoNet	MissouriNA	Average
Success (%)	36.8	71.8	65.0	57.9
Time (%)	19.2	57.7	51.0	42.6

The results of Table V show that both percentage values vary significantly between the different network topologies. Nevertheless, in all cases, the percentage of discarded solutions is always higher than the percentage of runtime spent by the heuristics. In the overall, almost 60% of the solutions were discarded at the cost of 42,6% of computational effort, showing that indeed the use of heuristics has improved the

overall computational efficiency of the proposed multi-start greedy randomized algorithm.

V. CONCLUSIONS

In this work, we have addressed the topology design of transparent optical networks aiming to maximize their resilience against critical node failures. We have proposed a multi-start greedy randomized algorithm resorting to a MILP based method, using row generation, to compute the critical nodes of each topology. The algorithm can be used both in the design of network topologies and in the upgrade of existing topologies.

We have run the proposed algorithm on three network topologies with publicly available information comparing the resiliency gap between the existing/upgraded topologies with the best topologies designed to maximize its resilience with the same fibre budget.

The results have shown that the resiliency gap of existing topologies is significantly large but network upgrades with $L' = 10\%L$ can already reduce significantly the resiliency gaps provided that such upgrades are aimed at maximizing the network resiliency to multiple node failures.

Finally, comparing the best topologies with the existing ones, the best topologies are characterised by a decrease of the number of nodes with the lowest and highest degrees and an increase of the number of nodes with degrees closer to the average node degree. This clearly shows that network topologies resilient to critical node failures tend to have more homogeneous degrees among all their nodes.

ACKNOWLEDGEMENTS

This work is based upon work from COST Action CA15127 ("Resilient communication services protecting end-user applications from disaster-based failures – RECODIS") supported by COST (European Cooperation in Science and Technology). First author was supported by FCT ("Fundação para a Ciência e a Tecnologia"), Portugal, through PhD grant with reference SFRH/BD/132650/2017. Second author was supported by FCT through project UID/EEA/50008/2013. Third author was supported by FCT through CIDMA within project UID/MAT/04106/2013.

REFERENCES

- [1] J. Rak, et al. *RECODIS: Resilient communication services protecting end-user applications from disaster-based failures*. in *proc. ICTON*, We.D1.4., 2016.
- [2] T. Gomes, et al. *A survey of strategies for communication networks to protect against large-scale natural disasters*. in *proc. RNDM*, 11–22, 2016.
- [3] M. Furdek, et al. *An overview of security challenges in communication networks*. in *proc. RNDM*, 43–50, 2016.
- [4] F. Barbosa, A. de Sousa, A. Agra. *The Design of Transparent Optical Networks Minimizing the Impact of Critical Nodes*. *Electronic Notes in Discrete Mathematics*, vol. 64, 165–174, 2018.
- [5] A. Agra, A. de Sousa, M. Doostmohammadi. *The Minimum Cost Design of Transparent Optical Networks Combining Grooming, Routing, and Wavelength Assignment*. *IEEE/ACM Trans. on Networking*, vol. 24, no. 6, 3702–3713, 2016.
- [6] A. Arulselvan, C. W. Commander, L. Eleftheriadou, P. M. Pardalos. *Detecting Critical Nodes in Sparse Graphs*, *Computers & Operations Research*, vol. 36, 2193–2200, 2009.
- [7] A. Veremyev, V. Boginski, E. Pasiliao. *Exact Identification of Critical Nodes in Sparse Networks via New Compact Formulations*. *Optimization Letters*, vol. 8, 1245–1259, 2014.
- [8] D. Santos, A. de Sousa, P. Monteiro. *Compact Models for Critical Node Detection in Telecommunication Networks*. *Electronic Notes in Discrete Mathematics*, vol. 64, 325–334, 2018.
- [9] M. Di Summa, A. Grosso, M. Locatelli. *Branch and cut algorithms for detecting critical nodes in undirected graphs*. *Comput. Optim. Appl.*, vol. 53, no. 3, 649–680, 2012.
- [10] A. Veremyev, O. Prokopyev, E. Pasiliao. *Critical nodes for distance-based connectivity and related problems in graphs*. *Networks*, vol. 66, no. 3, 170–195, 2015.
- [11] D.F. Rueda, E. Calle, J.L. Marzo. *Robustness Comparison of 15 Real Telecommunication Networks: Structural and Centrality Measurements*. *J. of Network and Systems Management*, vol. 25, no. 2, 269–289, 2017.
- [12] S. Neumayer, G. Zussman, R. Cohen, E. Modiano. *Assessing the Vulnerability of the Fiber Infrastructure to Disasters*. *IEEE/ACM Trans. on Networking*, vol. 19, no. 6, 1610–1623, Dec. 2011.
- [13] A. Zeng, W. Liu. *Enhancing network robustness against malicious attacks*. *Phys Rev E Stat Nonlin Soft Matter Phys*, vol. 85, no. 6, Pt 2, 2012.
- [14] A. Beygelzimer, G. Grinstein, R. Linsker, I. Rish. *Improving network robustness by edge modification*. *Physica A: Statistical Mechanics and its Applications*, vol. 357, no. 34, 593–612, 2005.
- [15] C. Natalino, A. Yayimli, L. Wosinska, M. Furdek. *Link addition framework for optical CDNs robust to targeted link cut attacks*. in *proc. RNDM*, 1–7, 2017.
- [16] F. Dikbiyik, M. Tornatore, B. Mukherjee. *Minimizing the Risk From Disaster Failures in Optical Backbone Networks*. *OSA Journal of Lightwave Technology*, vol. 32, no. 18, 3175–3183, 2014.
- [17] K.A. Sabeh, M. Tornatore, F. Dikbiyik. *Progressive network recovery in optical core networks*. in *proc. RNDM*, 106–111, 2015.
- [18] A. de Sousa, D. Mehta, D. Santos. *The Robust Node Selection Problem aiming to Minimize the Connectivity Impact of any Set of p Node Failures*. in *proc. DRCN*, Munich, 138–145, 2017.
- [19] G.T. Toussaint. *The relative neighbourhood graph of a finite planar set*. *Pattern Recognition*, vol. 12, no. 4, 261–268, 1980.
- [20] S. Orłowski, R. Wessaly, M. Piore, A. Tomaszewski. *SNDlib 1.0 — Survivable Network Design Library*. *Networks*, 55(3), 276–286, 2010.
- [21] S. Knight, H.X. Nguyen, N. Falkner, R. Bowden and M. Roughan. *The Internet Topology Zoo*. *IEEE J. on Selected Areas in Communications*, vol. 29, no. 9, 1765–1775, October 2013.
- [22] F. Rambach et al. *A multilayer cost model for metro/core networks*. *IEEE/OSA J. Optical Communications and Networking*, vol. 5, no. 3, 210–225, March 2013.