Embedded Network Design to Support Availability Differentiation

Abdulaziz Alashaikh · David Tipper ·
Teresa Gomes

Abstract The problem of how to provide, in a cost efficient manner, high levels of availability and service differentiation in communication networks was investigated in [1–3]. The strategy adopted was to embed in the physical layer topology a high availability set of links and nodes (termed the *spine*). The spine enables through protection, routing and cross layer mapping, the provisioning of differentiated classes of resilience with varying levels of end-to-end availability. Here we present an optimization model formulation of the spine design problem, considering link availability and the cost of upgrading link availability. The design problem seeks to minimize the cost while attaining a desired target flow availability. Extensive numerical results illustrate the benefits of modifying the availability of a subset of links of the network to implement quality of resilience classes.

Keywords crosslayer mapping · differentiated services · flow availability.

1 Introduction

Research on the resilience of communications networks that studies structural and operational aspects of the design of communications networks has been ongoing for decades [4]. The cost associated with providing high levels

Abdulaziz Alashaikh · David Tipper
Graduate Telecommunications and Networking Program
School of Computing and Information
University of Pittsburgh
Pittsburgh, PA 15260
E-mail: azizoozi@gmail.com . tipper@tele.pitt.edu

Teresa Gomes
Department of Electrical and Computer Engineering
University of Coimbra / INESC Coimbra
3030-290 Coimbra, Portugal
Email: teresa@deec.uc.pt

of resilience and the varying resilience needs of services has given rise to the problem of supporting resilience differentiation in communication networks [5]. The predominant approach to supporting multiple levels of resilience is based on assigning flows different protection and restoration schemes at a certain technology layer based on their Quality of Resilience (QoR) classes [6]. However, such a scenario-based approach can result in a resilience level below what is needed or inefficient utilization of network resources. Availability based routing strategies have appeared in the literature. A routing approach to manage the conflict that usually arises when seeking to ensure high availability while minimizing resource consumption was proposed in [7]. A routing algorithm to provide services with guaranteed differentiated availability, integrating protection schemes (dedicated and shared) was proposed in [8]. A computational model developed in [9] allows to exactly calculate the availability of a connection with shared backup path protection, in certain scenarios. The level of resilience required by some clients can be time-varying: their services may only require a stringent availability during certain time windows when some planned critical tasks are performed, while a lower availability may be acceptable during the remaining time periods. A framework to effectively address these time-differentiated resilience requirements was proposed in [10]. The need to maintain failure status for all flows and make real time routing decisions, is the main disadvantage of this approach.

In general, there are some limitations of the existing methods. First, the range and the spacing between availability classes for existing approaches are somewhat narrow. Both need to be enlarged to cover a broader range of classes. Second, using basic protection schemes such as 1+1, might be insufficient to achieve high availability for mission-critical services (*i.e.,* four to six 9's) [11]. For example, in [12], the 1+1 recovery scheme of the gold class is insufficient to support extremely high availability levels (*e.g.,* five or six 9's). Different ways to improve availability include increasing the level of dedicated protection (i.e., increasing the number mutually disjoint paths), reserving adequate shareable spare capacity to restore traffic from multiple simultaneous failures (*e.g.,* dual failure shared backup path protection [13]) and controlled flooding of traffic along dissemination graphs [14]. These approaches, however, lead to inefficient use of network resources and are also constrained by network diversity. Therefore, in some cases, it is necessary to augment the network by adding new links and possibly nodes to the network topology to support additional routes. Nonetheless, it is difficult to economically justify expanding a core network solely to improve its availability.

An alternative approach to achieving high availability is to increase the availability of network components. The authors of [15, 16] try to optimize network availability by improving the availability of a subset of physical links via shielding. In addition, Botton et al. [17] study a network design problem with a subset of edges that for a given cost can be upgraded to be more reliable. They show that to resort to a set of more reliable edges, as an alternative to pairs of edge-disjoint paths, may reduce resource utilization. These approaches, however, do not support resilience differentiation.
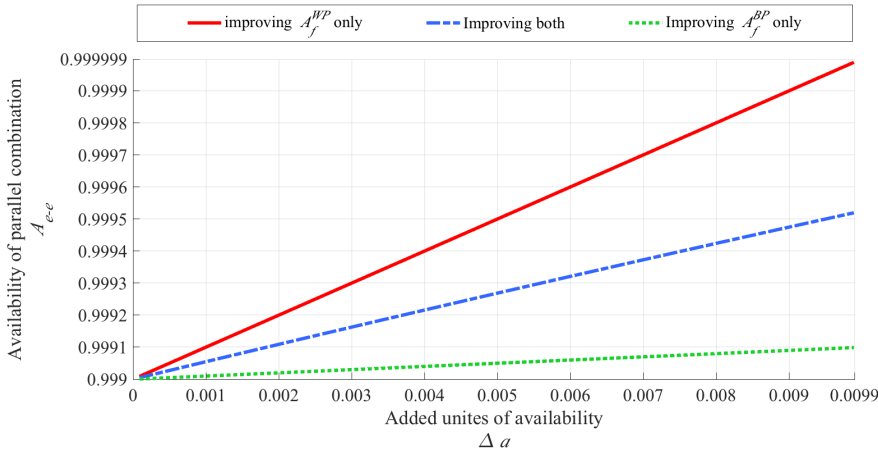
Fig. 1: Improving availability in a parallel configuration

The third limitation of existing methods is related to the application of these approaches to layered networks. With just a few exceptions, most of the existing approaches suffer from crosslayer mapping issues [18], as without full knowledge of the physical layer and the mappings between layers no hard guarantees on availability can be provided (*i.e.,* due to fault propagation).

Our approach to provide high and differentiated levels of availability stems from Brinbaum's importance measure. According to this measure, improving the component with the higher availability in a parallel configuration yields the best overall availability [19]. To illustrate this, consider a flow f routed over a working path (WP) and a disjoint backup path (BP) with availability $A_f^{WP} = 0.99$ and $A_f^{BP} = 0.90$, respectively. The end-to-end availability A_{e-e} of flow f is based on its A_f^{WP} and A_f^{BP}, and is calculated as a parallel configuration in which $A_{e-e} = 1 - (1 - A^{WP})(1 - A^{BP})$. Assume we want to strengthen one or both of the links by adding some availability units Δa, with options of $(A_f^{WP} + \Delta a)$, $(A_f^{BP} + \Delta a)$, or $(A_f^{WP} + \Delta a/2$ and $A_f^{BP} + \Delta a/2)$ to add these units. Figure 1 plots the overall end-to-end availability of flow f for the three options. It shows the best end-to-end availability A_{e-e} is obtained when improving A_f^{WP} only. It follows from this analysis that the use of high availability components in a working path in parallel with a protection path of lesser availability can lead to higher availability than would be obtained using homogeneous components. Adopting this concept, a network provider can allocate investment towards improving network reliability in an economic-efficient way via implementing a network spine.

Definition *The spine is a substructure with comparatively higher availability embedded into a network at the physical layer to improve the overall network availability without substantial modifications to the topology.*

The spine concept requires designing a network with heterogeneous link availabilities such that a substructure of the network has relatively larger availability values. The high availability substructure portion of the network is termed the *spine*. The spine is the foundation to provide differentiated classes of resilience, as it would connect nodes with traffic demands requiring a high level of availability. For example, the highest quality of resilience class traffic could be routed on the spine or use the spine as a backup path. The nodes, link interfaces and links on the network spine would have higher availability than the equipment that is not part of the spine. This provides *levels of availability differentiation at the physical level* which can be leveraged with restoration techniques, logical virtual network topology routing, cross layer mapping and other methods to further differentiate resilience classes and provide an extended range of availability guarantees. Although the spine could be any (connected) subgraph of the network, in our previous work [1–3] it was designed as a minimum spanning tree embedded at the physical layer, which was then used to route the connections of the higher availability QoR classes. Heuristics that use structural properties of graphs to calculate candidate spines were proposed, and a design performance study was carried out. Numerical results demonstrate that the spine approach effectively increases the range of network flows availability and allows high-availability values to be achieved that cannot be attained with standard protection configurations. In [2, 3], we assumed all links on the spine have the same availability (a_S) and similarly all links off the spine have the same availability (a_O) with ($a_O < a_S$ and $a_S = a_O + \Delta$). Our sensitivity analysis in [3] shows that modifying either value; the improvement step of availability Δ, or considering heterogeneous link availabilities, results in a change in the ranking of the best spines with respect to the availability metrics considered. In a follow on to our work, we showed that using spine-aware routing and crosslayer mapping, the spine concept provides levels of availability differentiation in multilayer networks with upper layer [20] or lower layer restoration [21]. In this paper, we revisit our proposal of the spine concept and formulate the spine design problem as a mixed integer linear programming (MILP) problem that considers the cost of upgrading link availability. In order to be able to evaluate the design of the spine focusing on cost and availability, other constraints were not considered. Hence, for simplicity we consider the uncapacitated network case and note that the our previous work [20, 21] showed only modest increases in the capacity are needed with the spine approach with the amount depending on the percentage of highest QoR class traffic.

The remainder of the paper is organized as follows. In Section 2, we provide some background on the spine and describe our model. In Section 3, we present the spine design optimization problem. In Section 4, we conduct a numerical study to evaluate our model and show a sample of the results. We conclude our paper in Section 5.

2 The Spine Model

We adopt an optimization model approach to determining the spine that involves taking into consideration factors such as different possible availability values to improve the links (and corresponding cost) and a target level of availability. The problem can be stated as follows: Given a physical network $\mathcal{G}_P = (\mathcal{V}_P, \mathcal{E}_P)$ with a set of nodes $\mathcal{V}_P$ and a set of physical links $\mathcal{E}_P$, a set $\mathcal{K}$ of improvement options for each link with associated cost c_{ij}^k; $(i,j) \in \mathcal{E}_P, k \in \mathcal{K}$ and a set of end-to-end connections/flows $(s,t) \in \mathcal{F}$ that need high availability, one seeks to determine the subnetwork, that forms the spine $\mathcal{G}_S = (\mathcal{V}_S, \mathcal{E}_S)$, where $\mathcal{G}_S \subset \mathcal{G}_P$, that minimizes the total cost while achieving an availability target. Before presenting the optimization problem the link availability and cost models are discussed.

2.1 Incremental Link Availability Model

For a given network, each link is assigned an initial link availability value a_{ij} based on its length, with longer links being less reliable. Specifically we use a distance-based link availability formula found in [22]. The link availability is calculated as $a_{ij} = a_{c_{ij}} \times a_{t_{ij}}$ where $a_{t_{ij}}$ is the product of cable-ends equipments (i.e., OXC, ROADM etc...), and $a_{c_{ij}}$ is the fiber cable availability that can be calculated from:

$$a_{c_{ij}} = 1 - \frac{MTTR}{MTBF} \tag{1}$$

$$MTBF_{hrs} = \frac{CC \times 365 \times 24}{\text{cable length}_{km}} \tag{2}$$

where CC is the cable cut rate, $MTBF$ and $MTTR$ are the mean time between failures and mean time to repair in hours, respectively. According to this model, cables which are the same length are expected to have the same estimated availability. However, in reality several other factors impact cable availability (*e.g.,* cable layout, physical protection, isolation, geographic route, proximity to maintenance centers, etc.) and result in differences in link availability despite having the same length. For example, aerial cables are less reliable then buried cables [23]. Furthermore, there are differences within the availability of buried cables depending on shielding, isolation and how the cable is buried (direct, microtrenching, ducted, etc.) [15]. Also, cables that traverse different terrains (metropolitan, rural, forest, etc), experience different rates of construction work and different weather conditions [24] which results in different failure rates. Lastly, it is worth nothing that cables have different proximity to highways, repair parts storage, or maintenance centers [25] which directly impacts the repair time.

In practice, it is expected that a network operator manages its network availability through a combination of different methods. However, operators

are constrained by the CAPEX/OPEX allocated for modifications, operations and maintenance, leaving the methods adopted subject to their efficiency (expected gain) and cost. Here, we study the scenario where each link in a network, can be purposely strengthened so that its MTBF is increased, for example by modifying the cabling (*e.g.,* burying an aerial cable) [26, 27] or adding physical protection [15, 16] or the MTTR is reduced by focused maintenance and repair efforts [26, 28, 29]. For each link, the possible options to increase their availability can be collected and each one will result in different levels of availability and cost. Specifically, if the link e, with end nodes i and j, has availability a_{ij}, using method k, the link availability can be augmented to a_{ij}^k with cost c_{ij}^k, whereas using method $k + 1$ that costs c_{ij}^{k+1}, availability is augmented to a_{ij}^{k+1}. We assume "K" possible availability values $(a_{ij}^k, k = 1, 2, ...K)$ with $(a_{ij}^1 = a_{ij})$. For each value k, the corresponding unavailability is decreased by ϵ, so that $u_{ij}^k = u_{ij}^{k-1} \cdot (1 - \epsilon)$, where $u_{ij}^k = 1 - a_{ij}^k$. Decreasing a link unavailability is akin to decreasing its expected downtime.

Note that, in reality the k different options might not have fixed downtime differences within nor across links. Here, we choose a fixed ϵ for illustration purpose. In real network planning scenarios, the number and type of alternatives for link availability augmentation will be different as this depends on several factors (*e.g.,* the terrain, cable type, the associated cost, regulations, etc). Nevertheless, we assume an identical set of possibilities of availability improvement across all the links, in order to simplify the model. The cost associated with each improvement step k is calculated by a cost function, $c_{ij}^k = f_c(a_{ij}^k, a_{ij}^1)$.

An accurate formula to calculate the cost of availability is difficult to obtain in practice. For this reason, many researchers resort to known mathematical models (*e.g.,* constant, linear, quadratic, etc) to relate cost to availability [30]. Also, note that the rate of return on investment decreases with increasing availability, as the cost to enhance a link availability becomes greater as availability increases. For example, bettering an already high availability link (*e.g.,* 0.999 to 0.9999) is typically more expensive than ameliorating a link with a modest availability (*e.g.,* 0.9 to 0.9009) by the same amount.

We consider the following cost functions, f_c's, to compute the cost of improving the link availability per unit of length. The cost function, f_{c1}, is a polynomial in the availability improvement $\Delta a_{ij}^k = a_{ij}^k - a_{ij}^1$.

$$f_{c1}(a_{ij}^k, a_{ij}^1) = \left(a_{ij}^k - a_{ij}^1\right)^\alpha \qquad , \ k > 2 \tag{3}$$

where α is a scaling parameter. This function informs that the greater the increase in availability, the greater the cost. The second cost function, f_{c2}, is a polynomial in the availability improvement $\Delta a_{ij}^k = a_{ij}^k - a_{ij}^1$ but also weighted by the unavailability of the link. Hence for equal Δa_{ij}^k, it augments the cost for the link with higher availability. This formula is very similar to f_2 in [31]

$$f_{c2}(a_{ij}^k, a_{ij}^1) = \left(\frac{a_{ij}^k - a_{ij}^1}{1 - a_{ij}^1}\right)^\alpha \qquad , \ k > 2 \tag{4}$$

The third cost function, f_{c3}, is derived from f_1 in [31] and conveys the idea that the effect on the availability of increased investment decreases exponentially.

$$f_{c3}(a_{ij}^k, a_{ij}^1) = -\ln\left[\frac{1 - a_{ij}^k}{1 - a_{ij}^1}\right] \qquad , \; k > 2 \tag{5}$$

The impact of link length on the respective availability upgrade cost is represented by the expression,

$$c_{ij}^k = f_{cl}(a_{ij}^k, a_{ij}^1) \times d_{ij} \quad l = 1, 2, 3 \tag{6}$$

where d_{ij} is link (i, j) length. Observe that the shorter the link the less the cost to improve the availability.

Figure 2 shows the Polska network topology and the availability options for three different links with $K = 7$. Each table in the figure shows the availability levels of a link and the corresponding cost for the different cost functions. The links availability was set to be in the interval [0.95,0.995], according to a linear variation of the edges' length (0.995 availability for the shortest edge and 0.95 for the longest edge). For each function, the obtained costs across all links are scaled between 1 and 100. The exponent α in (3) and (4) was set to 2 to impose quadratic growth of the cost. Here, $k = 1$ corresponds to the initial link availability. The case of $k = 2$ describes the transfer of maintenance resources between links, *i.e.,* increasing MTTR on some link(s) to decrease on other(s). Therefore, we set $u_{ij}^2 = u_{ij}^1 \cdot (1+\epsilon)$ and $c_{ij}^2 = -c_{ij}^3$. Thus the expected downtime of a link with $k = 2$ (*i.e.,* off the spine link) would increase and incur negative cost which would reduce the total cost C. It is assumed, as shown in the tables of Figure 2, that $f_{cz}(a_{i,j}^2, a_{ij}^1)$ ($z = 1, 2, 3$) is always the symmetrical value of $f_{cz}(a_{i,j}^3, a_{ij}^1)$, however, this is only strictly true for $z = 1, 2$. Furthermore, it is possible to improve the availability of other links (*i.e.,* in the spine) through the redistribution of maintenance and repair resources and taking advantage of the transfer of operational expenses of degraded links.

Figure 3 plots the cumulative distribution function (CDF) of the cost values for all links in the Polska network. Recall that the values obtained by the cost functions are normalized and placed on a scale of 1 to 100. One can see that f_{c1} has a smaller range of values. Further for f_{c1}, 75% of the cost values are below 20 for the Polska network whereas only 27% and 40% of these values are below 20 for f_{c2} and f_{c3}, respectively. In the other studied networks, an equivalent cost variation was observed.

3 The Spine Link Selection Design Problem

The design problem of the spine is to jointly select the set of arcs that define it and the corresponding options to increase their availability, in order to satisfy availability requirements, ensuring that the solution has minimum cost. In formulating the design problem we route all flows on the spine with fully link-disjoint backup paths. This ensures that all high QoR priority traffic supported by the spine can be given 1+1 dedicated protection. Note that, this

$e_{1\text{-}2}$	length=321 km				...
k	$a_{ij}^{\,k}$	$c_{ij}^{\,k}$			...
		f_{c1}	f_{c2}	f_{c3}	...
1	0.95551	0	0	0	...
2	0.93327	-20	-20	-15	...
3	0.97776	20	20	15	...
4	0.98888	43	52	34	...
5	0.99444	59	72	53	...
6	0.99722	67	84	71	...
7	0.99861	72	90	90	...

$e_{3\text{-}8}$	Length=259 km				...
k	$a_{ij}^{\,k}$	$c_{ij}^{\,k}$			...
		f_{c1}	f_{c2}	f_{c3}	...
1	0.96565	0	0	0	...
2	0.94848	-10	-15	-12	...
3	0.98283	10	15	12	...
4	0.99141	21	41	27	...
5	0.99571	29	57	42	...
6	0.99785	33	67	57	...
7	0.99893	35	72	72	...

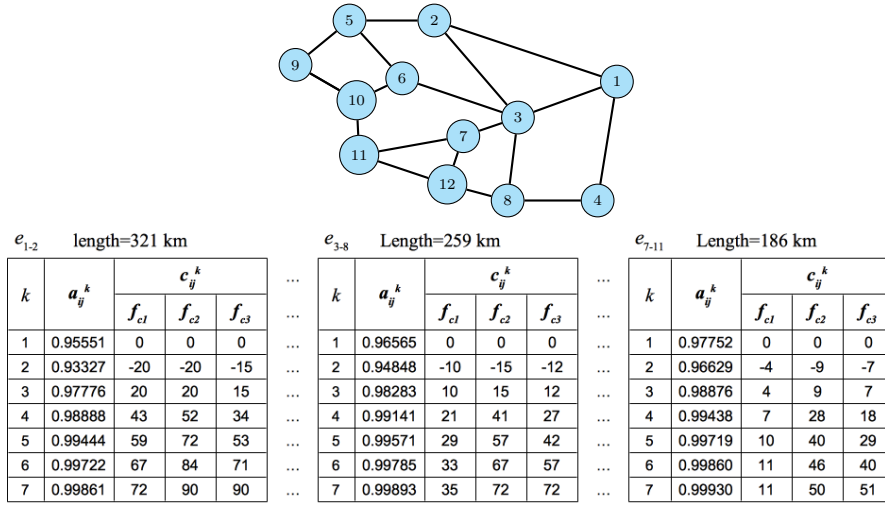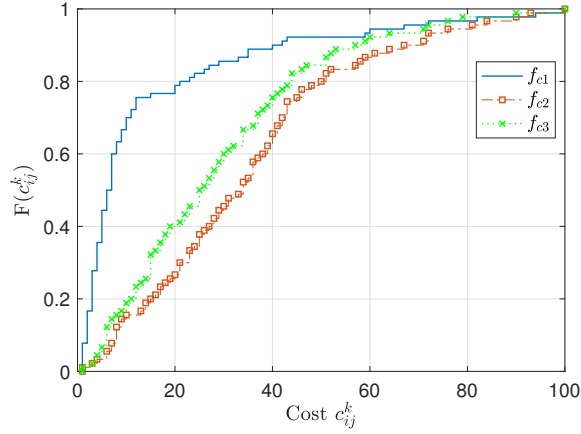| $e_{7\text{-}11}$	Length=186 km			
k	$a_{ij}^{\,k}$	$c_{ij}^{\,k}$		
		f_{c1}	f_{c2}	f_{c3}
1	0.97752	0	0	0
2	0.96629	-4	-9	-7
3	0.98876	4	9	7
4	0.99438	7	28	18
5	0.99719	10	40	29
6	0.99860	11	46	40
7	0.99930	11	50	51

Fig. 2: Polska network incremental link model.

Fig. 3: Polska network link improvement costs CDF.

also enables 1:N shared protection, however this topic is left for future work. We also assume that among all pairs of nodes there is at least one traffic flow of the QoR class of service which requires high availability (*e.g.,* a full mesh of demand of one unit between each node pair), so the spine will be a minimum spanning tree (MST). The high QoR class flow availability is constrained to be greater or equal to target availability values $(\widehat{a_{wp}}, \widehat{a_{bp}})$. The goal is to obtain the spine that guarantees each WP (in the spine) satisfies a minimum value of target availability $\widehat{a_{wp}}$, instead of maximizing the average availability. As

shown earlier, increasing the availability on the working path improves the end to end availability more effectively than improving all links at once. Finally, the objective of the design problem aims at minimizing the total costs of embedding the spine and improving flow availabilities to reach or exceed the target value.

3.1 Notation

Indices:

ij	represent a physical link by its two end-nodes, $i, j \in \mathcal{V}_P$, $(i, j) \in \mathcal{E}_P$.
st	represent a connection/flow between two end-nodes, $s, t \in \mathcal{V}_P$, $(s, t) \in \mathcal{F}$.

Parameters:

k	method of link improvement.
a_{ij}^k	availability of link (i, j) after applying improvement option k, with $a_{ij}^1 = a_{ij}$, (a_{ij}: initial link availability).
$\widehat{a_{wp}}(\widehat{a_{bp}})$	flow WP (BP) availability target.
$A_{\mathcal{S}}$ $(A_{\mathcal{S}}^{WP})$	average end-to-end flow (flow WP) availability.
c_{ij}^k	cost of improving link (i, j) using method k.
$H_{\mathcal{G}}$	total number of (undirected) links required by shortest (min-hop) path pairs.

Variables:

x_{ij}	a binary variable indicating whether link (i, j) is selected on the spine (x_{ij}=1) or not (x_{ij}=0).
x_{ij}^{st} (y_{ij}^{st})	a binary variable denoting whether physical link (i, j) is used for routing the WP (BP) of connection (s, t).
r_{ij}^k	a binary variable indicating if method k is used for link (i, j).
p_{ij}^{st} (q_{ij}^{st})	a continuous variable denoting link (i, j) unavailability given that it is on connection (s, t) WP (BP).

3.2 Optimization Model Formulation

Given the notation above, the spine design problem can be formulated as an optimization model as follows:

$$\textbf{Minimize} \qquad C = \sum_{ij} \sum_k r_{ij}^k \times c_{ij}^k \qquad (7)$$

s.t.

WP and BP computation:

$$\sum_{hj \in \mathcal{E}_P} x_{hj}^{st} - \sum_{ih \in \mathcal{E}_P} x_{ih}^{st} = \begin{cases} 1 & \text{if } h = s \\ -1 & \text{if } h = t \\ 0 & \text{otherwise} \end{cases} \qquad , \forall h \in \mathcal{V}_P, (s, t) \in \mathcal{F} \qquad (8)$$

$$\sum_{hj \in \mathcal{E}_P} y_{hj}^{st} - \sum_{ih \in \mathcal{E}_P} y_{ih}^{st} = \begin{cases} 1 & \text{if } h = s \\ -1 & \text{if } h = t \\ 0 & \text{otherwise} \end{cases} \qquad , \forall h \in \mathcal{V}_P (s, t) \in \mathcal{F} \qquad (9)$$

Loopless Routing:

$$x_{ij}^{st} + x_{ji}^{st} \leq 1, \qquad\qquad y_{ij}^{st} + y_{ji}^{st} \leq 1 \qquad\qquad ,\forall(i,j) \in \mathcal{E}_P, (s,t) \in \mathcal{F} \qquad (10)$$

$$\sum_{\substack{j \in \mathcal{V}_P \\ ij \in \mathcal{E}_P}} x_{ij}^{st} + \sum_{\substack{h \in \mathcal{V}_P \\ hi \in \mathcal{E}_P}} x_{hi}^{st} \leq 2 \qquad\qquad ,\forall i \in \mathcal{V}_P, (s,t) \in \mathcal{F} \qquad (11)$$

$$\sum_{\substack{j \in \mathcal{V}_P \\ ij \in \mathcal{E}_P}} y_{ij}^{st} + \sum_{\substack{h \in \mathcal{V}_P \\ hi \in \mathcal{E}_P}} y_{hi}^{st} \leq 2 \qquad\qquad ,\forall i \in \mathcal{V}_P, (s,t) \in \mathcal{F} \qquad (12)$$

Disjointness constraints:

$$x_{ij}^{st} + y_{ij}^{st} \leq 1 \qquad\qquad x_{ji}^{st} + y_{ji}^{st} \leq 1 \qquad\qquad ,\forall(i,j) \in \mathcal{E}_P, (s,t) \in \mathcal{F} \qquad (13)$$

Hop-count constraint:

$$H_{\mathcal{S}} = \sum_{ij \in \mathcal{E}_P} \sum_{st \in \mathcal{F}} (x_{ij}^{st} + y_{ij}^{st}) \leq \delta \times H_{\mathcal{G}}, \qquad\qquad \delta \geq 1 \qquad (14)$$

MST formation:

$$x_{ij} \geq x_{ij}^{st}, \qquad \forall(s,t) \in \mathcal{F} \qquad (15) \qquad x_{ij} = x_{ji}, \qquad \forall(i,j) \in \mathcal{E}_P \qquad (16)$$

$$\sum_{\substack{ij \in \mathcal{E}_P \\ i < j}} x_{ij} \leq |\mathcal{V}_P| - 1 \qquad (17)$$

Availability constraints:

$$\sum_{k} r_{ij}^{k} = 1, \qquad \forall(i,j) \in \mathcal{E}_P \qquad (18) \qquad r_{ij}^{k} = r_{ji}^{k} \qquad ,\forall(i,j) \in \mathcal{E}_P \qquad (19)$$

$$p_{ij}^{st} = x_{ij}^{st} \times \sum_{k} r_{ij}^{k} (1 - a_{ij}^{k}) \qquad\qquad ,\forall(i,j) \in \mathcal{E}_P, (s,t) \in \mathcal{F} \qquad (20)$$

$$q_{ij}^{st} = y_{ij}^{st} \times \sum_{k} r_{ij}^{k} (1 - a_{ij}^{k}) \qquad\qquad ,\forall(i,j) \in \mathcal{E}_P, (s,t) \in \mathcal{F} \qquad (21)$$

Flow availability targets:

$$A_{st}^{WP} = 1 - \sum_{ij \in \mathcal{E}_P} p_{ij}^{st} \geq \widehat{a_{wp}} \qquad\qquad (22)$$

$$A_{st}^{BP} = 1 - \sum_{ij \in \mathcal{E}_P} q_{ij}^{st} \geq \widehat{a_{bp}} \qquad\qquad (23)$$

Variables:

$$x_{ij}^{st}, y_{ij}^{st}, x_{ij}, r_{ij}^{k} \qquad \text{binary} \qquad (24) \qquad\qquad p_{ij}^{st}, q_{ij}^{st} \qquad \in [0,1] \qquad (25)$$

The essence of the formulation is the flow conservation constraints (8) and (9), which find primary and backup paths for all flows. A flow conservation constraint pushes a unit of demand along a path between the two end-nodes of a given flow. Constraint sets (10)-(12) ensure loop free routing. The set of constraints (13) guarantees fully link-disjoint primary and backup paths. The total number of links used in those path pairs $H_{\mathcal{S}}$, however, is limited by (14)

where δ is a scaling factor and $H_{\mathcal{G}}$ is the sum of the total number of hops of the shortest path pairs between all node-pairs in the network.

Each link used by a primary path of any flow is considered as an on-spine link. Constraint (15) ensures this by setting the spine link selector variable for a link, x_{ij}, to 1 if the link is used in a primary path of at least one flow. Due to the route of different (s,t) flows and equation (15), both x_{ij} and x_{ji} may take the value 1, but not necessarily both every time. Hence, because the network is undirected, constraint (16) is required. Then, constraint (17) limits the number of the (undirected) links selected for the spine to $|\mathcal{V}_P| - 1$ which is the number of links for a spanning tree. Next, constraint set (18) guarantees only one improvement method can considered for each link. Constraint (19) enforces the same improvement method in both directions of a link. Constraint sets (20) and (21) are used to associate a flow WP and BP unavailability to the unavailability of each link along the flow path. Variable p_{ij}^{st} or q_{ij}^{st} will have an unavailability value only if the WP or BP of flow (s,t) is routed through link (i,j). These two sets of constraints, transform the optimization problem into an integer nonlinear programming (INLP) model, because the product of two variables *i.e.,* x_{ij}^{st} with r_{ij}^{k} in (20) and y_{ij}^{st} with r_{ij}^{k} in (21). Note that, to compute a single path availability for a given flow, one can multiply the availability of the links along the path, but this results in a nonlinearity. Alternatively, we use the approximate version of the unavailability formula for a system connected in series, $(u^{st} \approx \sum_{ij} u_{ij}^{st})$. Hence, WP availability can be computed as $(1 - \sum_{ij} p_{ij}^{st})$. BP availability is computed in the same way. Constraints (22) and (23) guarantee the availability values of the AP and the BP of a flow are greater than the specified lower bounds, $\widehat{a_{wp}}$ and $\widehat{a_{bp}}$, respectively. Lastly, constraint sets (24) and (25) list binary and continuous variables.

To remove the nonlinearity of the INLP, we reformulate the availability constraints following the approach in [32]: constraints set (20) can be replaced with constraint sets eqs. (26) to (28). The three constraints provide the same function as (20). Similarly, constraints set (21) that computes BP unavailability can be replaced with the set of eqs. (29) to (31).

$$p_{ij}^{st} \leq x_{ij}^{st} \qquad (26) \qquad\qquad p_{ij}^{st} \leq \sum_{k} r_{ij}^{k} \, (1 - a_{ij}^{k}) \qquad (27)$$

$$p_{ij}^{st} \geq x_{ij}^{st} + \sum_{k} r_{ij}^{k} \, (1 - a_{ij}^{k}) - 1 \qquad\qquad \forall (i,j) \in \mathcal{E}_P, (s,t) \in \mathcal{F} \qquad (28)$$

$$q_{ij}^{st} \leq y_{ij}^{st} \qquad (29) \qquad\qquad q_{ij}^{st} \leq \sum_{k} r_{ij}^{k} \, (1 - a_{ij}^{k}) \qquad (30)$$

$$q_{ij}^{st} \geq y_{ij}^{st} + \sum_{k} r_{ij}^{k} \, (1 - a_{ij}^{k}) - 1 \qquad\qquad \forall (i,j) \in \mathcal{E}_P, (s,t) \in \mathcal{F} \qquad (31)$$

Although nonlinear constraints are avoided, the spine link selection design problem is NP-complete, since the optimization version of the problem of finding a min-sum disjoint path-pair with additional constraints is known to be

NP-complete [33–37]. Thus it is difficult to solve the problem for optimality. Instead, we can find a feasible solution within an optimality gap from a solution bound. There are a number of well known relaxation and heuristics methods for solving integer programming problems and iteratively trying to minimize the optimality gap. Hence one can expect solving the design problem efficiently for moderate sized networks and the chance of proving optimality becomes strongly size dependent.

4 Numerical Study

Here we present sample results for three network topologies often adopted in the literature [38], namely: Polska, Spain, and Italia14 networks. Table 1 summarizes the networks showing the number of nodes and links, the density ratio, and the diameter of each network.

G_P	Network	$|V_P|$	$|E_P|$	$\frac{|E_P|}{|V_P|}$	diameter
1	Polska [38]	12	18	1.50	4
3	Spain [39]	14	22	1.57	5
3	Italia14 [40]	14	29	2.07	3

Table 1: Test Networks.

First we examine the spine obtained from the optimization model and compare it to the spines obtained from heuristics in [2, 3], respectively, along with results from an exhaustive search evaluating all of the possible spanning trees. We consider the Polska network with $K = 2$, and two link availability values; $a_{ij}^1 = 0.99$ for off the spine and $a_{ij}^2 = 0.999$ for the on spine links. We solved the optimization problem with a minimum WP availability goal $\widehat{a_{wp}} = 0.995$ and the BP availability goal $\widehat{a_{bp}}$ of constraint (23) is relaxed. Then we calculated the average $A_{\mathcal{S}}^{WP}$ and $A_{\mathcal{S}}$ for the spine determined from the optimization problem and added them to the corresponding plots in [2, 3] that show the results of each metric for all possible spanning trees and the spines obtained from the heuristics given in [3], whose results are shown as red diamonds in each plot. Figure 4 shows the results for the spine obtained by the ILP as green circles along each line. We can see that the spine determined from the ILP has a high score with respect to $A_{\mathcal{S}}^{WP}$ and $A_{\mathcal{S}}$. The result shows that by ensuring a minimum flow WP availability on a spine, we can obtain a fairly good spine with high availability compared to an exhaustive evaluation of all possible spanning trees in a network.

Next we studied the more general case with the number of link types (*i.e.,* possible different availability values for each link) set to $K = 7$ for the networks of Table 1. We set δ in constraint (14) to 110%, allowing for a maximum of 10% increase in total resources over the resources required by shortest path-pairs, $H_{\mathcal{G}}$ (*i.e.,* $H_{\mathcal{S}} \leq 1.1 H_{\mathcal{G}}$). In the incremental link availability model $\epsilon =$
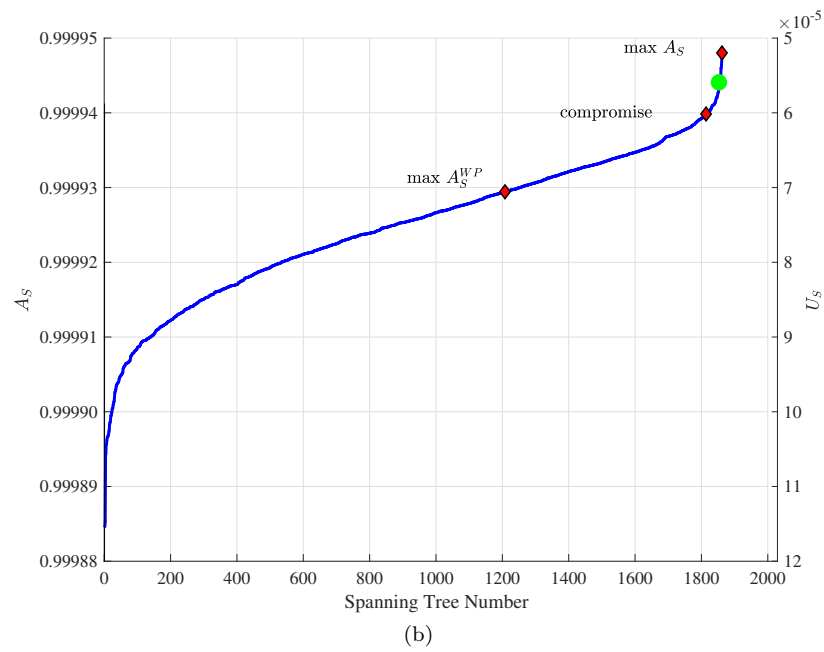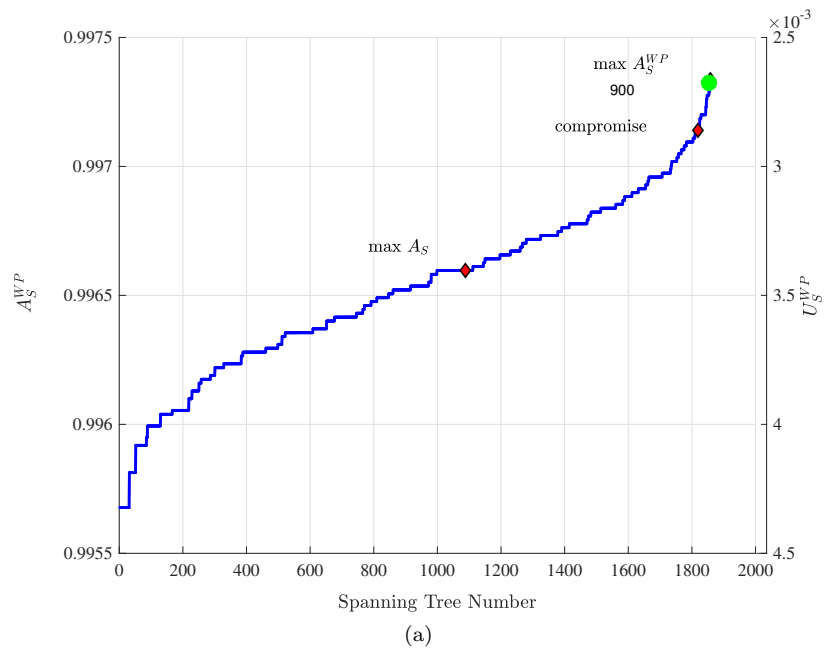
(a)

(b)

Fig. 4: Comparison of optimization based spine with heuristic spines and all spanning trees of the Polska Network

0.5 was used. The working path availability goal $\widehat{a_{wp}}$ in Equation (22), for the Polska network is set to $\widehat{a_{wp}} = \{0.99, 0.995, 0.996, 0.9964\}$, for the Spain network $\widehat{a_{wp}} = \{0.99, 0.996, 0.9964, 0.9967\}$, and for the Italia14 network $\widehat{a_{wp}} = \{0.99, 0.995, 0.996, 0.997\}$. At each each $\widehat{a_{wp}}$ setting, we solved the model for the three cost functions while relaxing the $\widehat{a_{bp}}$ constraint.

Figures 5, 7 and 9 show the spine layouts obtained for the Polska, Spain, and Italia14 networks, respectively, as the target $\widehat{a_{wp}}$ increases for all three cost functions. In the figures, the solid and dotted red lines are the links comprising the spine. The links marked with solid red lines are included in the spine in all the scenarios for a given cost function, whereas the dotted lines represent links that are selected for that case. Several interesting observations can be made from the numerical results. First, there is only a slight modification of the spine topology with the alteration of the target availability, $\widehat{a_{wp}}$, or of the cost function. Note, that increasing the target value $\widehat{a_{wp}}$ increases the total cost of the design and results in changes in the spine layout. In fact, for all three networks and cost functions the topology for the most stringent availability target $\widehat{a_{wp4}}$ differs from the topologies for the other target availability values. However, one can see that there is a persistent substructure that emerges in almost all the spines for a network regardless of the cost function and target availability values (*e.g.,* the star-like substructure rooted at node 3 in the Polska network (Figure 5(d)) and the clustered star-like substructure at nodes 11 and 7 in the Italia14 network (Figure 9(d)). Looking across all cost functions and target availability values, there are three distinct spine topologies for the Polska network, two topologies for the Spain network and five for the Italia14 network, which is the densest network.

Table 2 shows the graph theoretic structural properties of the spines, namely: eb_S the average edge betweenness centrality of the spine, ed_S the average edge degree of the spine, h_S the average shortest path on the spine, and di_S the diameter of the spine. The first row shows the corresponding measures for the full physical layer graph $\mathcal{G}_P$ with no spine. Observe that the spines tend to have a comparatively small edge betweenness eb_S and average shortest path h_S and a large edge degree ed_S. Only in a few cases does the corresponding measure match the minimum (or maximum) value found by generating all spanning trees for the network as given in [3]. Notwithstanding, these results follow suit with the findings of the heuristic algorithm (reported in [2, 3]) with respect to the spine that maximizes A_S^{WP} and provides evidence that the spines tend to have a star-like or cluster of stars rather than a ring-like layout. Although the layout of the spine can be imputed to the structural importance of the links and nodes, it is also delineated by the cost associated with the links and their availability, as well as by the constraint on the total number of hops.

Interestingly, it was observed that similar spines within and across cost functions are likely to have different availability and link types. This is illustrated by Figures 6, 8 and 10 where downtime per year for each link versus the link length and the link improvement method/type k selected for each link is displayed for the cases corresponding to Figures 5, 7 and 9. In the figures, each

cost function	scenario	eb_S	ed_S	h_S	di_S
	$\mathcal{G}_g$	0.12	6.33	2.14	4
f_{c1}	$\mathcal{G}_S, \widehat{a_{wp_1}}$	0.26	4.91	2.89	6
	$\mathcal{G}_S, \widehat{a_{wp_2}}$	0.26	4.91	2.89	6
	$\mathcal{G}_S, \widehat{a_{wp_3}}$	0.26	4.91	2.89	6
	$\mathcal{G}_S, \widehat{a_{wp_4}}$	0.24	5.45	2.68	5
f_{c2}	$\mathcal{G}_S, \widehat{a_{wp_1}}$	0.26	4.91	2.89	6
	$\mathcal{G}_S, \widehat{a_{wp_2}}$	0.26	4.91	2.89	6
	$\mathcal{G}_S, \widehat{a_{wp_3}}$	0.26	4.91	2.89	6
	$\mathcal{G}_S, \widehat{a_{wp_4}}$	0.24	5.45	2.68	5
f_{c3}	$\mathcal{G}_S, \widehat{a_{wp_1}}$	0.26	4.91	2.82	5
	$\mathcal{G}_S, \widehat{a_{wp_2}}$	0.26	4.91	2.82	5
	$\mathcal{G}_S, \widehat{a_{wp_3}}$	0.26	4.91	2.89	6
	$\mathcal{G}_S, \widehat{a_{wp_4}}$	0.24	5.45	2.68	5

(a) Polska: $|\mathcal{V}|=12$, $|\mathcal{E}|=18$, $\frac{|\mathcal{E}|}{|\mathcal{V}|}=1.5$

cost function	scenario	eb_S	ed_S	h_S	di_S
	$\mathcal{G}_g$	0.10	6.82	2.27	5
f_{c1}	$\mathcal{G}_S, \widehat{a_{wp_1}}$	0.26	4.46	3.38	7
	$\mathcal{G}_S, \widehat{a_{wp_2}}$	0.26	4.46	3.38	7
	$\mathcal{G}_S, \widehat{a_{wp_3}}$	0.26	4.46	3.38	7
	$\mathcal{G}_S, \widehat{a_{wp_4}}$	0.24	4.62	3.11	6
f_{c2}	$\mathcal{G}_S, \widehat{a_{wp_1}}$	0.25	4.62	3.21	7
	$\mathcal{G}_S, \widehat{a_{wp_2}}$	0.26	4.46	3.38	7
	$\mathcal{G}_S, \widehat{a_{wp_3}}$	0.26	4.46	3.38	7
	$\mathcal{G}_S, \widehat{a_{wp_4}}$	0.24	4.62	3.11	6
f_{c3}	$\mathcal{G}_S, \widehat{a_{wp_1}}$	0.25	4.62	3.21	7
	$\mathcal{G}_S, \widehat{a_{wp_2}}$	0.26	4.46	3.38	7
	$\mathcal{G}_S, \widehat{a_{wp_3}}$	0.26	4.46	3.38	7
	$\mathcal{G}_S, \widehat{a_{wp_4}}$	0.24	4.62	3.11	6

(b) Spain: $|\mathcal{V}|=14$, $|\mathcal{E}|=22$, $\frac{|\mathcal{E}|}{|\mathcal{V}|}=1.57$

cost function	scenario	eb_S	ed_S	h_S	di_S
	$\mathcal{G}_g$	0.06	10.3	1.87	3
f_{c1}	$\mathcal{G}_S, \widehat{a_{wp_1}}$	0.23	5.38	2.97	6
	$\mathcal{G}_S, \widehat{a_{wp_2}}$	0.23	5.38	2.97	6
	$\mathcal{G}_S, \widehat{a_{wp_3}}$	0.22	5.54	2.86	6
	$\mathcal{G}_S, \widehat{a_{wp_4}}$	0.21	5.85	2.79	6
f_{c2}	$\mathcal{G}_S, \widehat{a_{wp_1}}$	0.21	6.00	2.71	5
	$\mathcal{G}_S, \widehat{a_{wp_2}}$	0.21	6.00	2.71	5
	$\mathcal{G}_S, \widehat{a_{wp_3}}$	0.21	6.00	2.71	5
	$\mathcal{G}_S, \widehat{a_{wp_4}}$	0.21	5.85	2.79	6
f_{c3}	$\mathcal{G}_S, \widehat{a_{wp_1}}$	0.21	6.00	2.71	5
	$\mathcal{G}_S, \widehat{a_{wp_2}}$	0.21	6.00	2.71	5
	$\mathcal{G}_S, \widehat{a_{wp_3}}$	0.21	6.00	2.71	5
	$\mathcal{G}_S, \widehat{a_{wp_4}}$	0.21	6.31	2.70	5

(c) Italia14: $|\mathcal{V}|=14$, $|\mathcal{E}|=29$, $\frac{|\mathcal{E}|}{|\mathcal{V}|}=2.07$
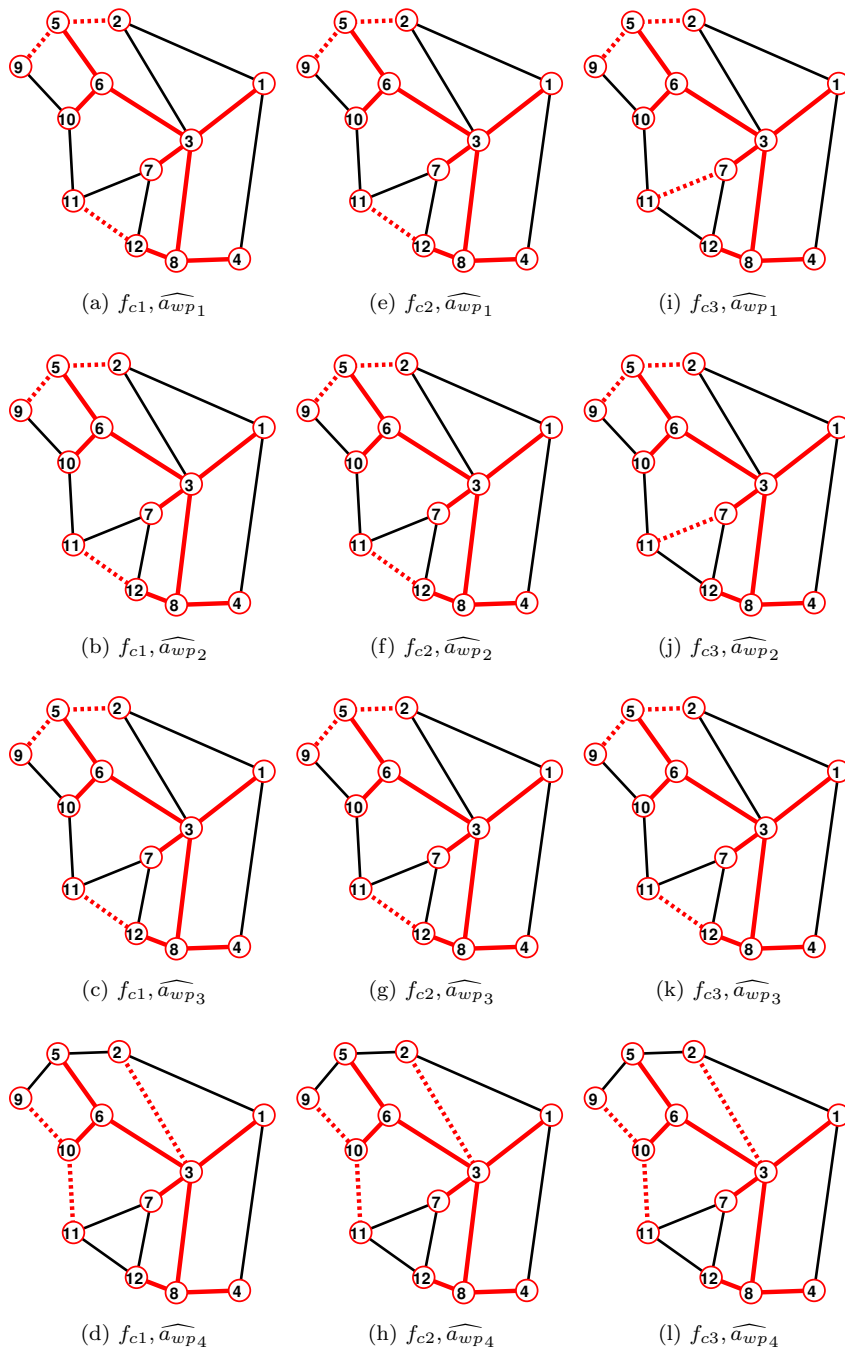
Table 2: Structural properties of the spines.

(a) $f_{c1}, \widehat{a_{wp_1}}$

(e) $f_{c2}, \widehat{a_{wp_1}}$

(i) $f_{c3}, \widehat{a_{wp_1}}$

(b) $f_{c1}, \widehat{a_{wp_2}}$

(f) $f_{c2}, \widehat{a_{wp_2}}$

(j) $f_{c3}, \widehat{a_{wp_2}}$

(c) $f_{c1}, \widehat{a_{wp_3}}$

(g) $f_{c2}, \widehat{a_{wp_3}}$

(k) $f_{c3}, \widehat{a_{wp_3}}$

(d) $f_{c1}, \widehat{a_{wp_4}}$

(h) $f_{c2}, \widehat{a_{wp_4}}$

(l) $f_{c3}, \widehat{a_{wp_4}}$

Fig. 5: Spines obtained for the Polska network using the three cost functions. Solid red line indicates a link has been selected to be on spine in all scenarios and dotted red lines the additional links on the spine.
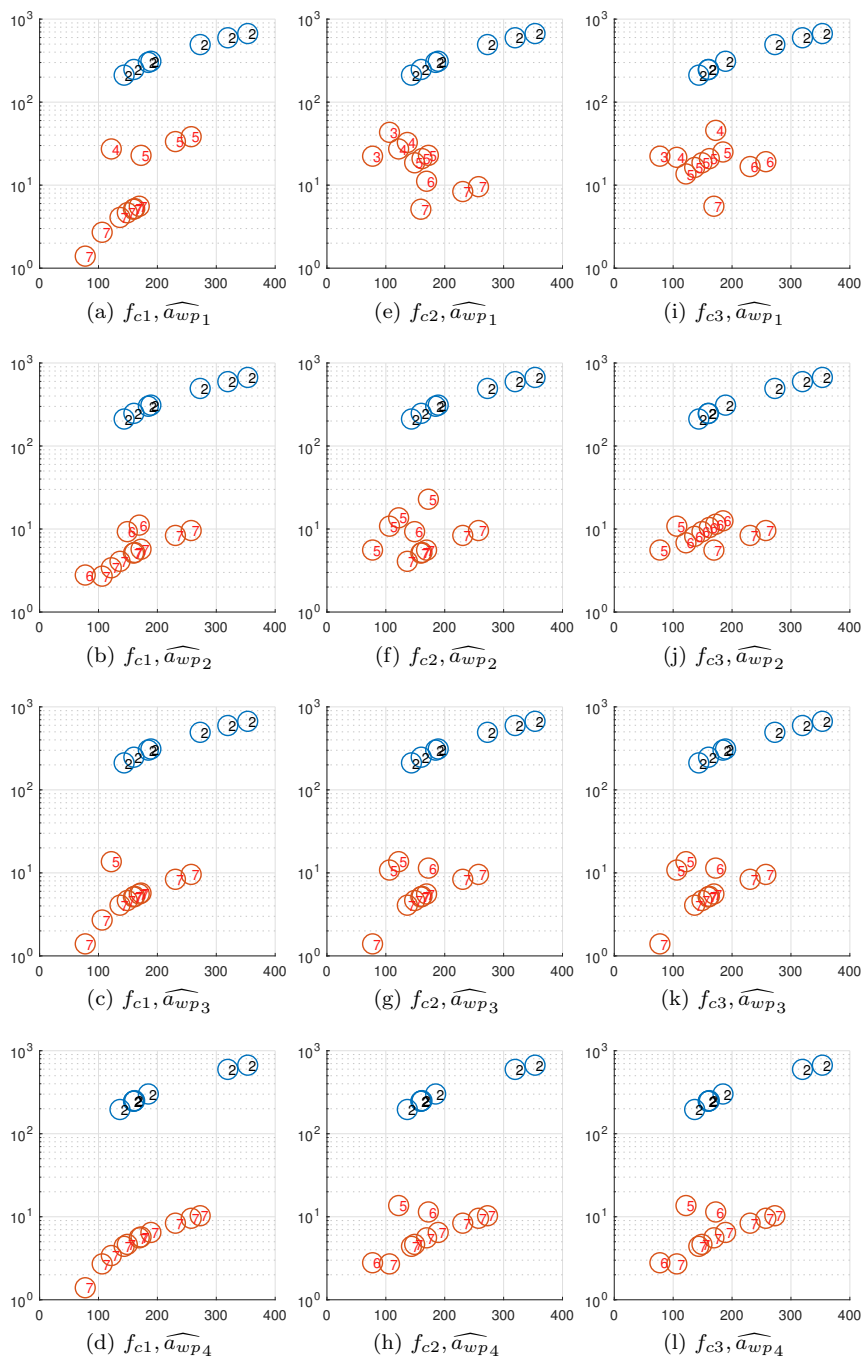
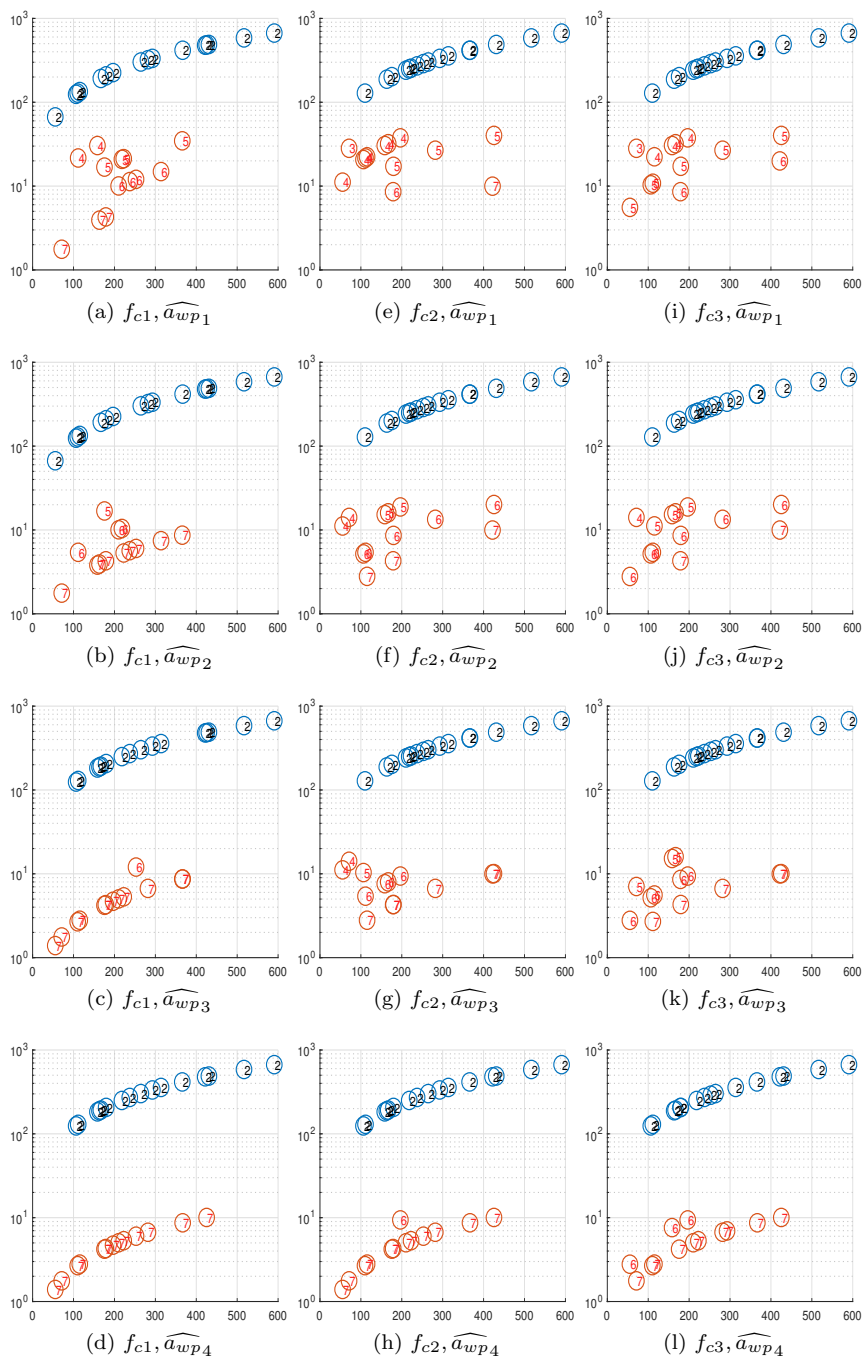Fig. 6: The corresponding link downtime/year and versus link length for the spines obtained for the Polska network.
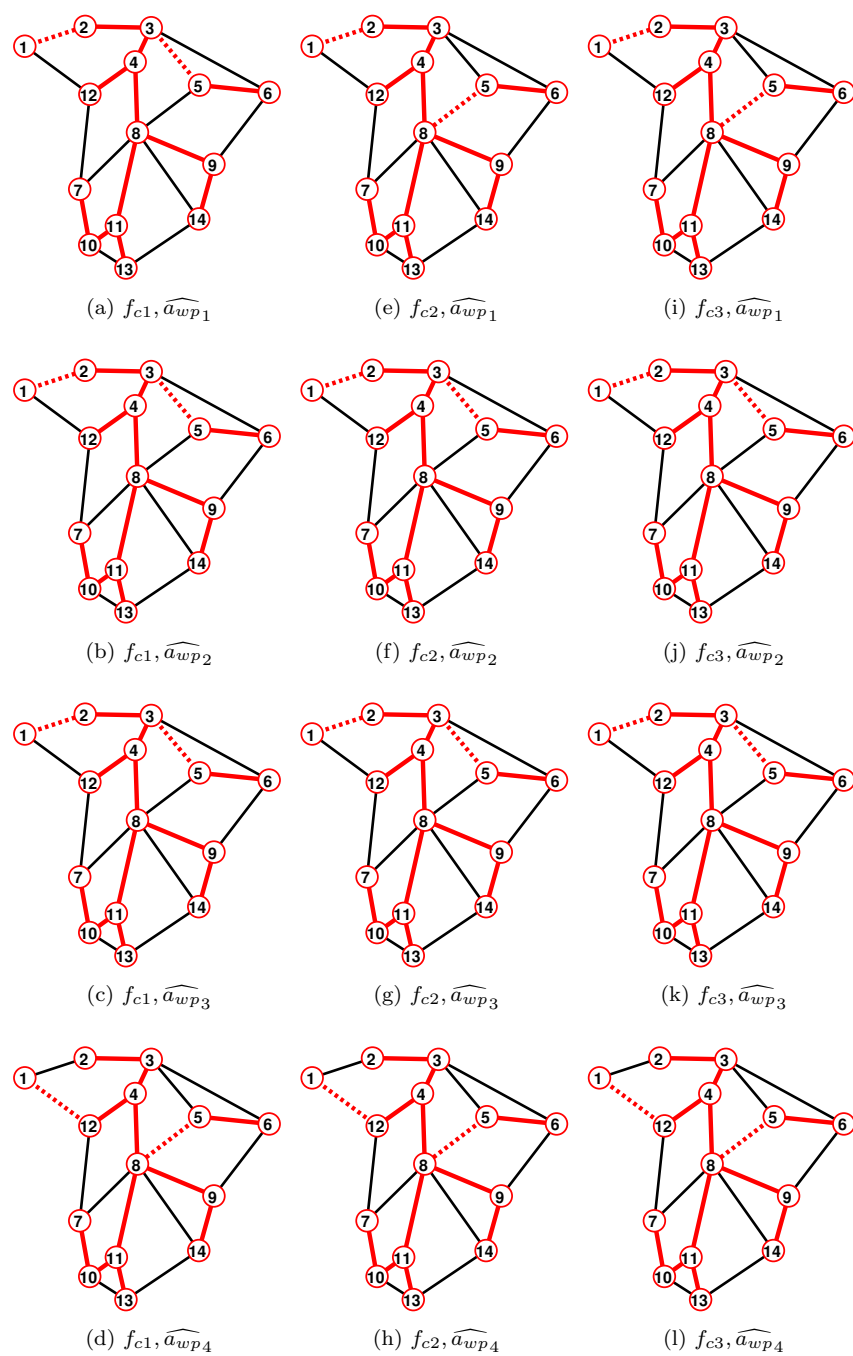
Fig. 7: Spines obtained for the Spain network using the three cost functions. Solid red line indicates a link has been selected to be on spine in all scenarios and dotted red lines the additional links on the spine.

Fig. 8: The corresponding link downtime/year and versus link length for the spines obtained for the Spain network.

(a) $f_{c1}, \widehat{a_{wp_1}}$

(e) $f_{c2}, \widehat{a_{wp_1}}$

(i) $f_{c3}, \widehat{a_{wp_1}}$

(b) $f_{c1}, \widehat{a_{wp_2}}$

(f) $f_{c2}, \widehat{a_{wp_2}}$

(j) $f_{c3}, \widehat{a_{wp_2}}$

(c) $f_{c1}, \widehat{a_{wp_3}}$

(g) $f_{c2}, \widehat{a_{wp_3}}$

(k) $f_{c3}, \widehat{a_{wp_3}}$

(d) $f_{c1}, \widehat{a_{wp_4}}$

(h) $f_{c2}, \widehat{a_{wp_4}}$

(l) $f_{c3}, \widehat{a_{wp_4}}$
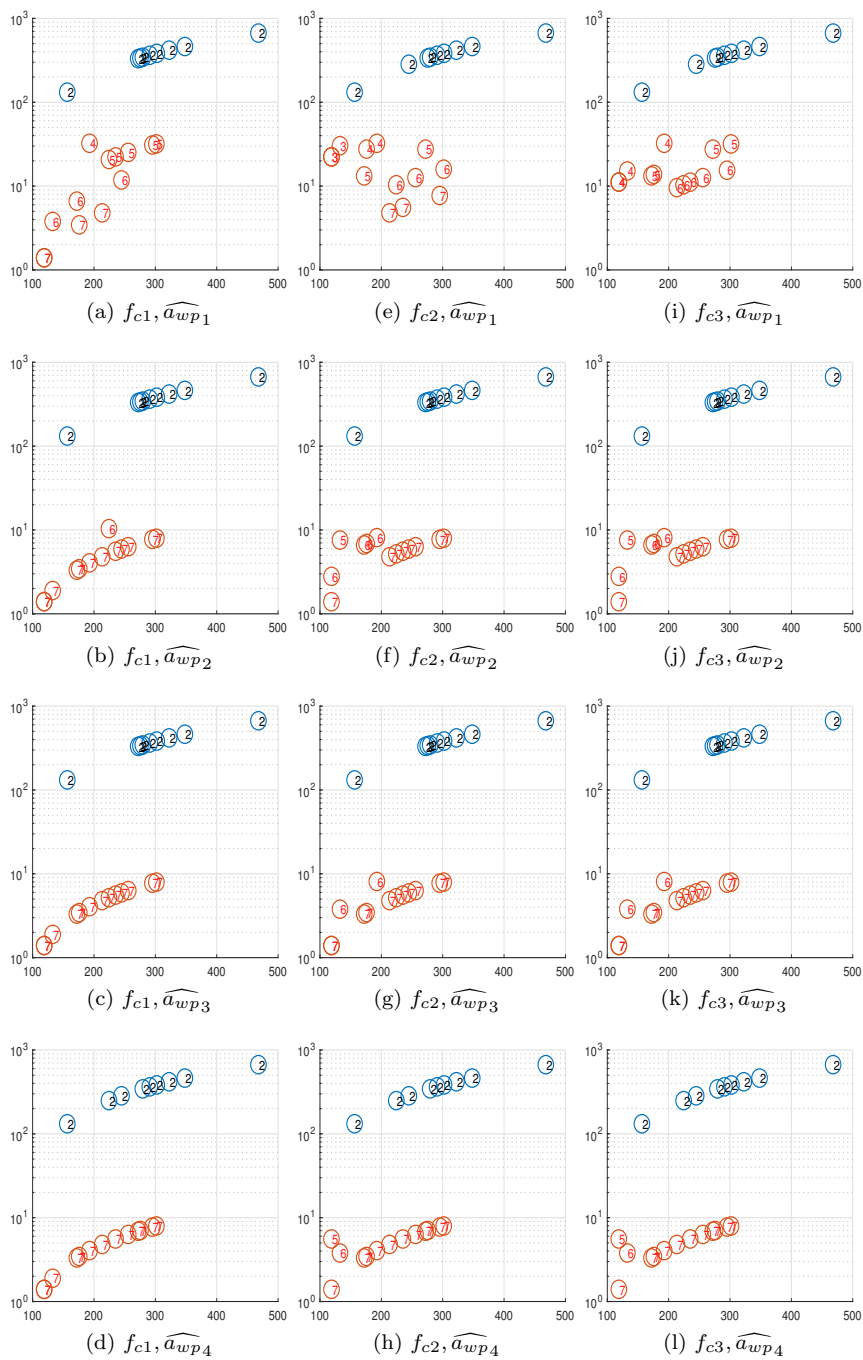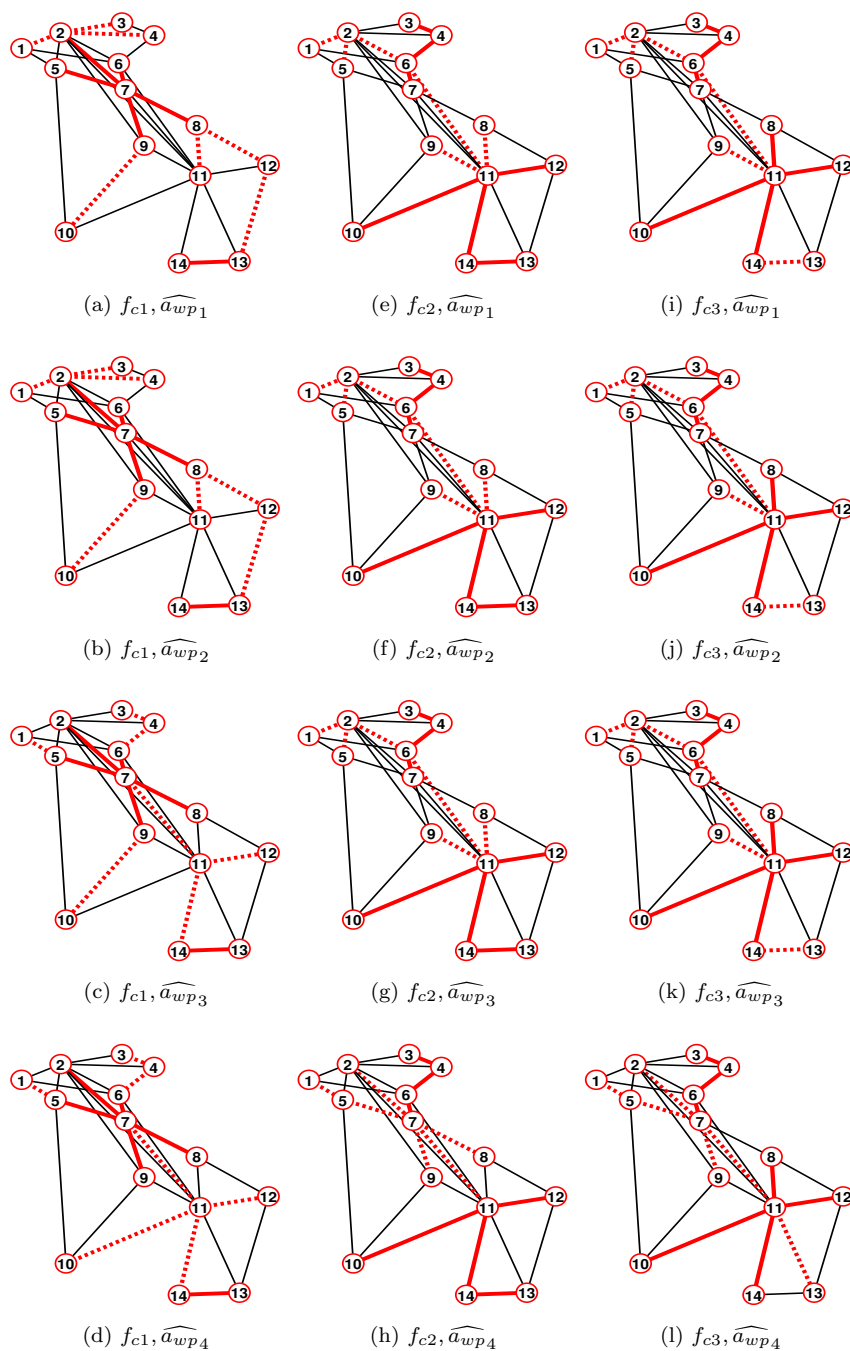
Fig. 9: Spines obtained for the Italia14 network using the three cost functions. Solid red line indicates a link has been selected to be on spine in all scenarios of the same cost function and dotted red lines the additional links on the spine.

Fig. 10: The corresponding link downtime/year and versus link length for the spines obtained for the Italia14 network.

circle represents a link and the number inside the circle is the improvement method/type k. The red circles identify the links that make up the spine and the blue ones correspond to the links off the spine. For example, note that the Polska network spines obtained with $\widehat{a_{wp4}} = 0.9964$ for cost functions f_{c1} and f_{c2} have the same layout, as displayed in Figures 5d and 5h, but the corresponding chosen link improvement methods k are distinct as seen in Figures 6d and 6h. However, comparing the Polska network spines with $\widehat{a_{wp4}} = 0.9964$ for f_{c2} and f_{c3} the spine layout and the selected methods are identical, whereas, they were significantly different with $\widehat{a_{wp1}}$ as shown in Figures 5e and 5i. It can also be observed that with a fixed cost function, as the WP availability target $\widehat{a_{wp}}$ increases, distinct link improvement methods k can be chosen. This is visible in Figures 5a to 5c, where the first three spines obtained for the Polska network for cost function f_{c1}, have the identical topology but differ in the selected improvement options, with the increase of the WP availability target $\widehat{a_{wp}}$. The downtime and availability assignment shown in Figures 6a to 6c illustrates this analysis. Initially, shorter links (*i.e.,* with higher availability and lower improvement cost) are favored as a spine link, thus exploiting existing heterogeneity. As the availability target increases, expensive links are selected to meet the more stringent requirement. For example, consider how the spine layout changes from the initial one of Figure 5a with $\widehat{a_{wp1}}$, to that of Figure 5d in order to achieve $\widehat{a_{wp4}}$. Note that link (2,3) (that is, the third longest link) is selected to be on the spine despite its high cost. Also Table 2 shows that the spine for $\widehat{a_{wp4}}$ has better structural measures (*e.g.,* smaller di_S) than the spine for $\widehat{a_{wp1}}$. The results for Spain and Italia14 networks lead to similar conclusions. Observe that the off spine links are chosen to be of type $k = 2$ which are less reliable, thus generating additional budget resources to upgrade the spine links from $k = 1$ to superior quality links (*i.e.,* $k > 2$).

We used AMPL/Gurobi on a NEOS server to solve our models to generate the spines of Figures 5, 7 and 9. We were able to get optimal solutions with 0% optimality gap for all cases. The corresponding optimization computational solution times are given in Figure 11. From the figure one can verify the solution time increases with the network size. Moreover, the cost function f_{c1} is the most computationally efficient to determine a spine.

In addition to the structural properties, we compared the average expected flow downtime $d_S = (1 - A_S) \times 8760\ hrs/year$ and average expected WP downtime $d^{WP} = (1 - A_S^{WP}) \times 8760\ hrs/year$ for the different scenarios and cost functions. We also include the corresponding downtime of an equivalent network with no spine and considering link-disjoint path pairs. When no spine is considered, all links are improvded identically (*i.e.,* using the same k method) and the total cost is calculated accordingly. Figure 12 shows the average expected WP and end-to-end flow downtimes for the test networks considering the different cost functions and $\widehat{a_{wp}}$'s. The results are shown for three cases: no spine (dotted line) which is the baseline case; then the case where the MTTR is relaxable for the off spine links and the third that forbids relaxing the MTTR for the off spine links (i.e., $k = 2$ is not allowed). From the figures we can make several observations. First, for all networks, the downtime values
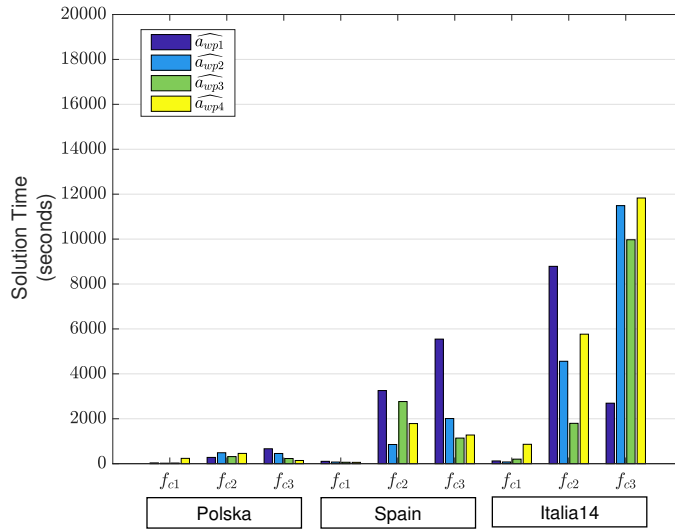
Fig. 11: Optimization computational solution times.

ameliorate over the no spine model, when using cost functions f_{c1} and f_{c2}, albeit the improvement is less remarkable for f_{c2}. Second, only in the case of the Italia14 network did cost function f_{c3} ensure a steady superiority over the no spine model. Nevertheless, Italia14 network, the one with the densest topology, attains a lower downtime than the no spine model, independent of the cost function adopted. Comparing relaxable and non-relaxable MTTR cases shows a significant saving in cost when relaxing the off spine links MTTR.

Keep in mind that the spine concept is intended to create different levels of availability and also satisfy the most severe availability requirements. Figure 13 shows the expected downtime for each path type (WP, BP and link-disjoint path pair) for the spines of Figures 5, 7 and 9. The downtime results are represented for each scenario as a box plot. The upper and lower edges of each box represent the third and first quartile of the values, respectively, the middle bar (in red) represents the median, and the upper and lower bars represent the maximum and minimum downtime values across all paths, respectively. Note that, even for the spine with the lowest cost (*i.e.,* $\widehat{a_{wp_1}}$ and relaxable $MTTR$), there are three different levels of availability classes resulting from using only one protection scheme. The lower availability class can be given an unprotected path with availability similar to a backup path, with larger expected downtime. Then, the middle class is routed on an unprotected path on the spine which achieves shorter expected downtime compared to the lower class. The higher class is routed on the spine and protected by a link-disjoint backup path, and its expected downtime is minimal.

Within each scenario, the range of availability for the middle class is upper bounded by the target availability, $\widehat{a_{wp}}$, as shown by the maximum downtime
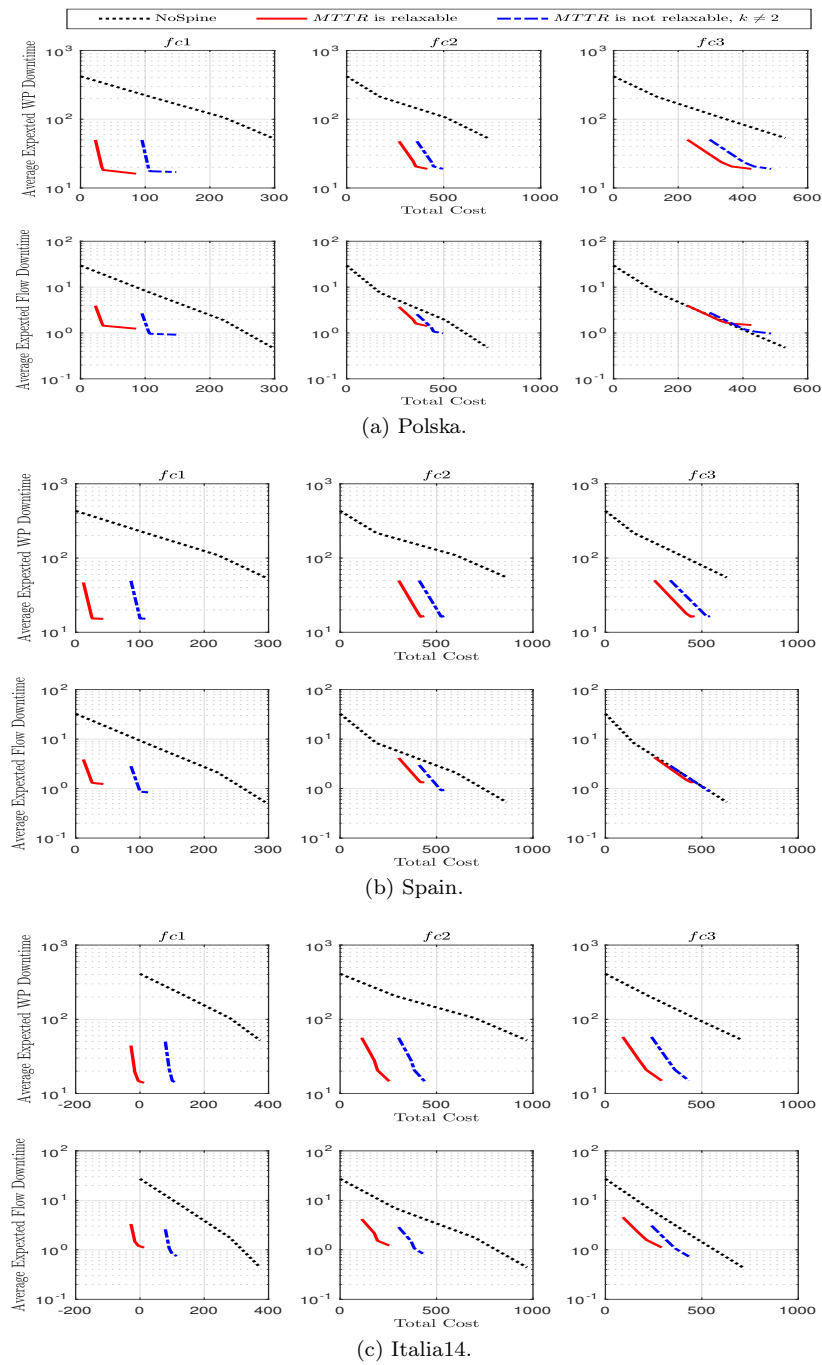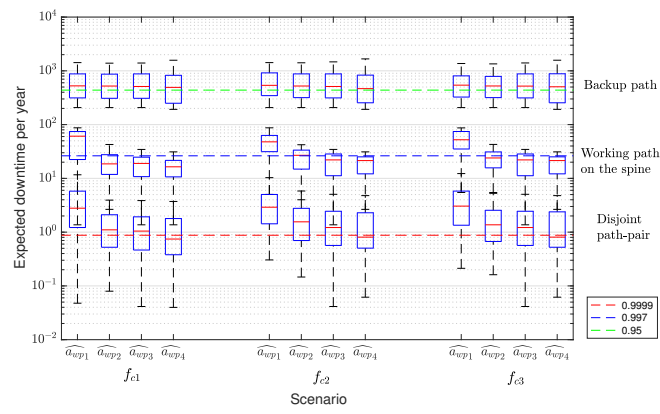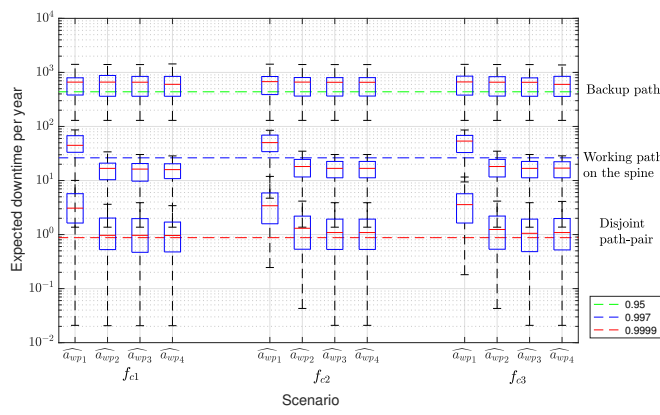
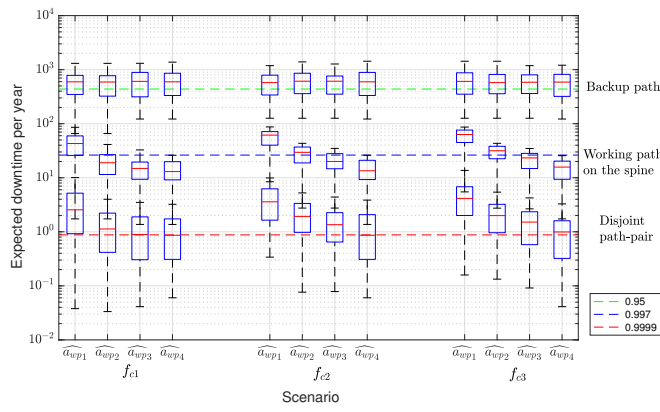Fig. 12: Average expected WP and end-to-end flow downtime/year versus cost

(a) Polska.

(b) Spain.

(c) Italia14.

Fig. 13: The range of expected downtimes in the spines

bar. The variation within that range, however, is attributed to the layout of the spine, and in general to the structure of the network, as some connections typically have longer paths on the spine. We expect that a spine with very long diameter to have a wider range of flows downtime. One also can see from the graph that the target availability also controls the downtime of the higher class since the WP of this class is routed on the spine, and its downtime decreases as the target availability increases. The maximum flow downtime within this class is close to the median downtime value compared to the minimum which means that the range of the flows downtime is moderately narrow above the median. This can also be attributed to the structure of the network but mainly is due to the WP target availability as the end-to-end availability of a path-pair is lower bounded by the availability of the highest path availability. For the lower class, the expected downtime is independent from the WP target availability, and it maintains a similar range of downtimes across the different scenarios. The spacing between each level of availability is mainly determined by the range of link availabilities (initial and improved) and the WP target availability.

5 Conclusions

The spine concept of embedding a subgraph structure with higher availability in a network together with protection mechanisms targeting to improve the overall end-to-end availability was revisited. The spine based approach was shown to have the potential to improve the network availability in a more efficient fashion compared to improving the availability of all network components in a homogeneous fashion. An optimization model formulation of the spine design problem, considering link availability and the cost of upgrading link availability, was provided. The design problem explores the heterogeneity of existing link availability and the ability to upgrade links to attain a desired flow availability while minimizing the total cost. Our results confirm the efficiency of the spine model in terms of average flow availability and its potential benefit over the shortest path model with no spine. This efficiency, however, depends primarily on network density and link improvement cost distribution. In general the spine provides larger differences in the range of availability values to quality of resilience classes resulting in less over engineering of the network to meet the most stringent availability requirements. Lastly, the proposed design problem can be solved efficiently only for moderate sized networks. Developing heuristic algorithms and scalable techniques are necessary to tackle larger instances of the problem.

Acknowledgment

References

1. D. Tipper, "Resilient network design: Challenges and future directions," *Telecommunication Systems*, vol. 56, no. 1, pp. 5–16, 2014.
2. T. Gomes, D. Tipper, and A. Alashaikh, "A novel approach for ensuring high end-to-end availability: The spine concept," in *10th International Conference Design of Reliable Communication Networks (DRCN)*, pp. 1–8, IEEE, 2014.
3. A. Alashaikh, T. Gomes, and D. Tipper, "The spine concept for improving network availability," *Computer Networks*, vol. 82, pp. 4–19, May 2015.
4. J. Rak, M. Pickavet, K. S. Trivedi, J. A. Lopez, A. M. C. A. Koster, J. P. G. Sterbenz, E. K. Çetinkaya, T. Gomes, M. Gunkel, K. Walkowiak, and D. Staessens, "Future research directions in design of reliable communication systems," *Telecommunication Systems*, vol. 60, no. 4, pp. 423–450, 2015.
5. P. Cholda, A. Mykkeltveit, B. Helvik, O. Wittner, and A. Jajszczyk, "A Survey Of Resilience Differentiaion Frameworks in Communication Networks," *IEEE Communications Surveys*, vol. 9, no. 4, pp. 32–55, 2007.
6. H. Zhang and A. Durresi, "Differentiated multi-layer survivability in IP/WDM networks," *IEEE/IFIP Network Operations and Management Symposium, NOMS 2002*, pp. 681–694, 2002.
7. B. Kantarci, H. T. Mouftah, and S. Oktug, "Arranging Shareability Dynamically for the Availability-Constrained Design of Optical Transport Networks," *Proceedings - IEEE Symposium on Computers and Communications*, pp. 68–73, 2008.
8. M. Tornatore, D. Lucerna, B. Mukherjee, and A. Pattavina, "Multilayer Protection with Availability Guarantees in Optical WDM Networks," *Journal of Network and Systems Management*, vol. 20, pp. 34–55, Sept. 2012.
9. L. Zhou, M. Held, A. Member, and U. Sennhauser, "Connection Availability Analysis of Shared Backup Path-Protected Mesh Networks," *Journal of Lightwave Technology*, vol. 25, no. 5, pp. 1111–1119, 2007.
10. M. Xia, M. Tornatore, S. Sevilla, L. Shi, C. U. Martel, and B. Mukherjee, "A Novel SLA Framework for Time-Differentiated Resilience in Optical Mesh Networks," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 3, no. 4, pp. 312–322, 2011.

11. D. A. Garbin, J. E. Knepley, F. Park, D. South, and F. Church, "Design and Analysis of High Availability Networks," in *IEEE Conference on Technologies for Homeland Security*, pp. 1–6, 2009.

12. S. Verbrugge, D. Colle, P. Demeester, R. Huelsermann, and M. Jaeger, "General availability model for multilayer transport networks," in *Design of Reliable Communication Networks, 2005. DRCN 2005. 5th IEEE International Workshop on*, 2005.

13. V.Y.Liu and D. Tipper, "Spare capacity allocation using shared backup path protection for dual link failures," *Computer Communications*, vol. 36, no. 6, pp. 666–677, 2013.

14. A. Babay, E. Wagner, M. Dinitz and Yair Amir, "Timely, Reliable, and Cost-Effective Internet Transport Service Using Dissemination Graphs," *IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, pp. 1-12, Atlanta, GA, June, 2017.

15. J. Zhang, E. Modiano, and D. Hay, "Enhancing network robustness via shielding," in *11th International Conference on the Design of Reliable Communication Networks (DRCN)*, pp. 17–24, March 2015.

16. J. Zhang, E. Modiano, and D. Hay, "Enhancing Network Robustness via Shielding," *IEEE/ACM Transactions on Networking*, vol. 25, pp. 2209–2222, August 2017.

17. Q. Botton, B. Fortz, and L. Gouveia, "On the hop-constrained survivable network design problem with reliable edges," *Computers & Operations Research*, vol. 64, pp. 159–167, 2015.

18. P. Pacharintanakul and D. Tipper, "Crosslayer survivable mapping in overlay-IP-WDM networks," in *Proceedings of the 2009 7th International Workshop on the Design of Reliable Communication Networks, DRCN 2009*, pp. 168–174, IEEE, 2009.

19. M. Rausand and R. Hoyland, *System Reliability Theory: Models, Statistical Methods and Applications*. Wiley Interscience, 2ed ed., 2003.

20. A. Alashaikh, D. Tipper, and T. Gomes, "Exploring the Logical Layer to Support Differentiated Resilience Classes in Multilayer Networks," *Annals of Telecommunications*, pp. 1–17, 2017.

21. A. Alashaikh, D. Tipper, and T. Gomes, "Supporting differentiated resilience classes in multilayer networks," in *12th International Conference on the Design of Reliable Communication Networks (DRCN)*, pp. 31–38, IEEE, mar 2016.

22. J.-P. Vasseur, M. Pickavet, and P. Demeester, *Network Recovery: Protection and Restoration of Optical, SONET-SDH, IP, and MPLS*. Morgan Kaufmann Publishers, Elsevier, 2004.

23. W. D. Grover, "Failure Impacts, Survivability Principles, and Measures of Survivability," in *Mesh-based Survivable Networks: Options and Strategies for Optical, MPLS, SONET and ATM Networking* (W. D. Grover, ed.), ch. 3, pp. 103–172, Prentice Hall PTR, 2003.

24. D. Crawford, "Fiber optic cable dig-ups: Causes and cures," 1993.

25. A. P. Snow, *A Reliability Assessment of the Public Switched Telephone Network Infrastructure*. PhD thesis, School o f Information Sciences, Pitts-

burgh, PA, USA, 1997. AAI9812414.

26. I. Rados, T. Sunaric, and P. Turalija, "Suggestions for Availability Improvement of Optical Cables," in *Circuits and Systems for Communications, 2002. Proceedings. ICCSC '02. 1st IEEE International Conference on*, pp. 234–239, IEEE, 2002.

27. H.-Y. Chang, I.-K. Chen, and W.-C. Chiang, "A Simple Network-Engineering Approach for Improving Connection Availability in WDM Networks," *2011 IEEE Workshops of International Conference on Advanced Information Networking and Applications*, pp. 805–807, Mar. 2011.

28. H.-Y. Chang and P.-C. Wang, "Upgrading service availability of optical networks: A labor force perspective," *International Journal of Communication Systems*, vol. 31, pp. 1–15, June 2018.

29. E. Marcus and H. Stern, *Blueprints for High Availability*. Wiley Publishing, 2nd ed., 2003.

30. T. Nagae and H. Wakabayashi, "Differences in Network Reliability Improvement by Several Importance Indices," *Transportation Research Procedia*, vol. 10, pp. 155–165, 2015.

31. U. Franke, "Optimal IT service availability: Shorter outages, or fewer?," *IEEE Transactions on Network and Service Management*, vol. 1, no. 9, pp. 22–33, 2012.

32. L. Liberti, S. Cafieri, and F. Tarissan, "Reformulations in Mathematical Programming: A Computational Approach," in Foundations of Computational Intelligence Volume 3: Global Optimization, pp. 153–234, Springer Berlin Heidelberg, 2009.

33. A. Pašić, P. Babarczi and A. Kőrösi, "Diversity coding-based survivable routing with QoS and differential delay bounds," in Optical Switching and Networking, vol. 23, pp. 118–128, 2017.

34. R. M. Karp, "Reducibility among combinatorial problems," in *50 Years of Integer Programming 1958-2008: From the Early Years to the State-of-the-Art*, pp. 219–241, Springer-Verlag, 2010.

35. J. Q. Hu, "Diverse routing in optical mesh networks," *IEEE Transactions on Communications*, vol. 51, pp. 489–494, March 2003.

36. C.-L. Li, S. Thomas McCormick, and D. Simchi-Levi, "Finding disjoint paths with different path-costs: Complexity and algorithms," *Networks*, vol. 22, no. 7, pp. 653–667, 1992.

37. A. Beshir and F. Kuipers, "Variants of the Min-Sum Link-Disjoint Paths Problem," in *16th Annual Symposium on Communications and Vehicular Technology in the Benelux (SCVT 2009)*, IEEE/SCVT, 2009.

38. S. Orlowski, M. Pióro, A. Tomaszewski, and R. Wessäly, "SNDlib 1.0–Survivable Network Design library," *Networks*, vol. 55, no. 3, pp. 276–286, 2010.

39. R. Martínez, R. Casellas, R. Vilalta, and R. Muñoz, "GMPLS / PCE-Controlled Multi-Flow Optical Transponders in Elastic Optical Networks [Invited]," *Journal of Optical Communications and Networking*, vol. 7, no. 11, pp. 71–80, 2015.

40. D. Colle, S. De Maesschalck, C. Develder, P. Van Heuven, A. Groebbens, J. Cheyns, I. Lievens, M. Pickavet, P. Lagasse, and P. Demeester, "Data-centric optical networks and their survivability," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 1, pp. 6–20, 2002.