BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS
DEPT. OF TELECOMMUNICATIONS AND MEDIA INFORMATICS
DOCTORAL SCHOOL OF ELECTRICAL ENGINEERING

INSTANTANEOUS FAILURE RECOVERY AND LOCALIZATION IN TRANSPORT NETWORKS

Alija Pašić
M.Sc. in Electrical Engineering

Summary of the Ph.D. Dissertation

Supervised by

Péter Babarczi Ph.D.
MTA-BME Future Internet Research Group
High Speed Networks Laboratory
Department of Telecommunications and Media Informatics
Budapest University of Technology and Economics

Budapest, Hungary
2018

1 Introduction

Communication networks are considered to be topmost critical infrastructures in today's information society. Owing to the rise of novel mission-critical applications (e.g. telesurgery, stock market), high reliability and the delay requirements gained on importance and became crucial. However, satisfying both requirements i.e., high reliability and low latency at the same time in a bandwidth-efficient manner is unquestionably one of the most challenging tasks for service providers in transport networks.

In addition to these challenges, the nature of transport networks has changed. The optical transport networks shifted from opaque to transparent optical (or all-optical) networks [17] to reach higher data rates and to avoid the time and energy consuming O/E/O (optical/electrical/optical) conversion. On the other hand, new technologies like SDN (Software Defined Networking) emerged and introduced an entirely new network management concept decoupling the network control and forwarding functions. SDN both abstracts the infrastructure for applications and network services, and makes the network control fully and directly programmable. It is not surprising that the new failure recovery and localization methods have to consider the needs and requirements of these new technologies and trends (all-optical networks and SDN).

The most important target of the Internet carriers is to meet the service requirements defined in the Service Level Agreement (SLA) with the subscribers in their backbones, i.e., to provide high Quality of Service (QoS). *In other words providing and maintaining a reliable network (i.e., providing high QoS) has utmost importance for network providers, because business customers are willing to pay significantly more for a reliable service [12]. This means that the providers have to model the network correctly (in hindsight of QoS) and choose the proper failure mitigation techniques (proper failure management techniques), while keeping the CAPEX (CAPital EXpenditures) and OPEX (OPErating EXpenditures) as low as possible.*

However, it is a well known fact that transferring user's data along a single active (or *working*) path in these new transport networks with throughput in order/magnitude of Tbytes per second [19] can not be sufficient to fulfill the service availability requirements in the presence of various network outages and unexpected failure events. This means in such transport networks even the shortest disruptions lead to huge amount of data loss i.e., huge drop in QoS, which is unacceptable. To avoid this, *instantaneous recovery* is required in today's transport networks.

Definition 1.1. *The requirement of* instantaneous recovery *is fulfilled if the recovery time is less then 50ms ($t_R \leq 50ms$) [16].*

In order to understand and investigate the recovery time (t_R), which is a critical (key) parameter of any protection scheme, we have to understand the recovery process (GMPLS [14]) of transport networks. The main steps are the following:

1. (failure localization time) t_l is defined as the time to locate the failure at a responsible network entity,
2. (failure notification time) t_n is the failure notification time, where the responsible network entity notifies all necessary switching nodes (via control plane signaling) which perform protection switching,
3. (failure correlation time) t_c is defined as the time period between the time instant when the switching node receives all the notifications, and the time instant when the failure is identified,
4. (decision time) t_d is the time during which the switching nodes select the necessary switching actions (protection paths),
5. (switching time) t_s is defined as the time for setting up the protection path i.e., configuration of the nodes, Optical cross-connects (OXC) etc. This step itself can take several tens of milliseconds [36, 37]).

$$t_R = t_l + t_n + t_c + t_d + t_s. \tag{1}$$

This means when planning the recovery process our main goal is to keep the recovery time under 50 ms, in order to ensure seamless operation even when a failure occurs. However, instantaneous recovery is not the only important factor for a protection approach. Besides instantaneous recovery, the robustness of the approaches can be crucial.

Definition 1.2. *A protection approach is* ***robust*** *if during the recovery process no control plane actions (messaging) are required and no OXC reconfiguration occurs [J1].*

In the first part of my dissertation I propose new methods with low resource requirements that ensure *instantaneous recovery in a robust manner* while being able to satisfy *delay constraint even after a failure* occurs. I compare the new methods with $1+1$ and other state-of-the-art protection approaches. In the second part of the dissertation I focus on the failure localization. I propose a new all-optical local failure localization framework, that enables fast restoration even for the shared protection approaches, reducing the failure recovery time significantly. Note that due to the nature of the shared protections, optical cross-connect (OXC) reconfiguration is always needed. This means *robust recovery can not be achieved* i.e., the simplicity level of $1 + 1$ can not be maintained. Since the OXC reconfiguration itself can take up to several tens of milliseconds depending on the technology [36, 37] i.e., instantaneous recovery can not be guaranteed, even if the failure localization and notification time takes only milliseconds.

2 Research Objectives

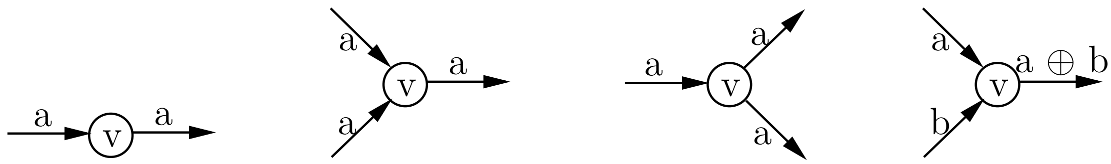
The objective of this dissertation is to propose novel methods for transport networks, which in addition to providing instantaneous recovery, are able to satisfy given delay constraints even after the failure occurs, with low resource consumption. In other words these new methods provide instantaneous recovery in a capacity efficient and robust manner, taking into account the delay requirements and the new technology trends (all optical networks, SDN) and their limitations.

The main technological constraints are the following:

– The network flows can not be divided into arbitrary many parts i.e., flow division is supported in a limited manner [J1].

– Inside the network only simple operations can be implemented [J1] like the XOR (all-optical XOR already can be implemented [20]). In Figure 1 the basic node capabilities are shown i.e., forwarding, merging, splitting, and network coding.

The protection approaches introduced in this dissertation reduce the trade-off and provide an appealing alternative in such way that the positive properties of $1 + 1$ are preserved (i.e., robustness and instantaneous recovery is provided) and in addition the



(a) Forwarding (b) Merging (c) Splitting (d) Coding

Figure 1: The basic node capabilities i.e. node roles. Each node v in the networks can be classified into one of this roles.

capacity allocation overhead is reduced. Note that although single link failures are the most common ones (more than 70% [13] of all failures affecting links) the backbone networks should be able to support connections with different level of availability (i.e., QoS level). In particular, in my dissertation, I propose new protection and failure localization frameworks, which not only consider the new trends in transport networks, but are also able to satisfy the new strict multi-constraint requirements. Particularly in the first part of my dissertation, a new, easily deployable and extremely resource efficient single link failure resilient method called Survivable Routing with Diversity Coding (SRDC) is introduced. The method is easily deployable, since the *coding itself can be done in the source and destination nodes, i.e., there is no need to upgrade network nodes.* Several subproblems of SRDC were investigated depending on capacity and node capability constraints in the network. For multiple failures the method called Survivable Routing with Network Coding (SRNC) was introduced.

Several works are dealing with Quality-of-Service (QoS) routing and differential delay (DD) aware survivable routing in transport networks; however all of the methods are designed for disjoint paths only. I generalize these results for diversity coding based survivable routing using Directed Acyclic Graphs (DAG). Hence, I define the before- and after-failure delay of a DAG, and investigate the effect of QoS routing and DD-aware routing delay constraints on these optimal survivable routing structures, ensuring instantaneous recovery.

In the second part of the dissertation the focus is on failure localization. I propose a new all-optical local failure monitoring framework that enables ultra fast (although not instantaneous) restoration even for the shared protection approaches, by reducing the failure recovery time significantly. I introduced a new concept called forbidden link-pairs which generalizes several monitoring trail based problems.

3 Methodology

The theoretical results and their properties in the dissertation were *obtained and assessed through analytical methods*, using the results of graph theory and computational complexity theory.

I implemented the proposed algorithms in C++ using the LEMON [3] network optimization library. *Extensive simulations were conducted* in order to validate the proposed methods and to compare the novel methods with approaches proposed in the literature. To obtain an optimal solution for the NP-complete problems I have followed the well accepted approach in the literature, i.e., I have formulated the problems as an Integer Linear Program. To solve the ILPs I have used the commercially available solvers GUROBI [2] and CPLEX [1].

4 New Results

Besides low latency the reliability i.e., the high availability of the telecommunication networks is one of the most important issues for the operators. In order to be able to investigate this subject, we have to model the network properly and reliably. Transport networks are usually represented in a layered manner [16]. In my dissertation the physical layer which consists of fibers and optical cross-connects is represented with a directed graph $G = (V, E)$. The nodes represent the OXC-s where the connection can enter or leave the networks. The set of edges E represents the bidirectional fiber connections between the OXCs (i.e., Optical Multiplex Section).

For the failure modeling the method called Shared Risk Link Group (SRLG) list ($\mathcal{F}$) is used. With the SRLGs we are able to handle dependent multiple failures. The SRLG are capable of expressing the statistical dependencies between failures (e.g., links, nodes). For example if we know that two seemingly unrelated links share a tunnel or a conduit on a given section, then these two links belong to the same SRLG since a single failure event like a tunnel fire (e.g. Baltimore tunnel fire) can disrupt both links. Note that an extensive part of the dissertation focuses on providing robust and instantaneous recovery methods for a very practical case i.e., single link failures scenario (the SRLG list contains only all the single link failures) which is the most common case (more than 70% [13] of failures are single link failures). Nonetheless several algorithms were presented for the general failures scenario (i.e., arbitrary SRLG list) in order to complete the study, and to be able to provide different QoS levels.

4.1 Instantaneous Failure Recovery in Transport Networks

In transport networks with throughput in order of Tbytes per second [19] even the shortest disruptions lead to huge amount of data loss i.e., huge drop in QoS, which is unacceptable. Nonetheless, due to huge distances transport networks are highly vulnerable to physical link failures and protection approaches have to be implemented [16]. In the protection of transport networks, three important aspects have to be taken into account. The combination of these three aspects determines the effectiveness or ineffectiveness of the protection approach.

The three aspects are: recovery time, complexity (both computational complexity and the complexity of the protected failure scenarios), and the resource consumption. Naturally, the question arises: Which one of the three aspects is the most important? To answer this question, we have to take a look on the practice.

In today's transport networks the most widely spread protection approach is the so called dedicated $1+1$ protection approach, owing to its simplicity and ultra fast recovery time. The dedicated $1+1$ protection gives the fastest restoration against single link failures, where the failure localization and notification time is zero, and the time setting up the protection path is assumed to be a few milliseconds. The process requires no control plane signaling, as each node can decide whether a failure occurred or not in a distributed fashion. Thus the $1+1$ protection provides *robust and instantaneous recovery*. In order to keep the complexity and restoration time low, this property has to be fulfilled. However, as by $1+1$ the data is sent parallel on both disjoint paths, fast recovery is provided to the connections for the price of *excessive amount of protection resources*. Besides that, $1+1$ optimizes only for one cost parameter (e.g. bandwidth or delay) and can not consider multiple constraints at the same time [7]. As the *new stricter QoS requirements desire multiple constraints*

$1+1$ has to be replaced in the long term.

Of course there are protection approaches like the shared protection approaches [10, 9], which have a much better resource consumption. By shared protection schemes a single protection path (also denoted as P-LP) can be used to protect several (disjoint) working paths i.e., can be shared. This means the user data is carried on the working paths only and the shared backup path is only used when a failure occurs affecting one of those (protected) working paths. However, after the failure has occurred, extensive signaling is required between the nodes of the path or the segments affected by the failure to restore the connection, leading to long recovery time for the price of efficient resource utilization. In other words, shared protection schemes yield better resource efficiency, however are not able to provide *robust and instantaneous recovery* and since instantaneous recovery is a must in transport networks, shared protection approaches are not yet an alternative.

Of course the $1+1$ is not the only *dedicated protection* approach. Previously reported studies on dedicated protection [7] have manipulated various assumptions on the node capability for the incoming and outgoing flows at each node to improve bandwidth efficiency, such as switching and merging, bifurcated or non-bifurcated, and with or without network coding.

Definition 4.1. *When* network coding *[24] is applied, instead of simple forwarding, the switches are able to perform algebraic operations on their incoming packets to construct the outgoing encoded packet.*

The General Dedicated Protection (GDP) [J1] framework was introduced for the dynamic routing and includes routing and network coding based protection approaches, too. With network coding the GDP is able to ensure instantaneous recovery with the minimum resource allocation possible for an arbitrary failures scenario, and is a *theoretical lower bound* for all dedicated protection approaches ensuring instantaneous recovery. However, this method can not be implemented in practice since the GDP with network coding (i.e., *Lower bound* algorithm) is very complex from the network equipment and management point of view, and dealing with arbitrary number of subflows is not possible in nowadays networks.

To resolve the practical implementation issue, I introduce a practical version of GDP, in which the flows are divided into exactly two parts [J1].

It is important to highlight that the solution of all methods with network coding consists of two parts. The *survivable routing* $R = (V^R, E^R, f)$ which contains the nodes and the edges with the given flow values of the solution. And the *robust configuration* $\mathcal{C}$, which includes all the information needed for the network coding and OXC configuration.

Note that if *survivable routing* $R = (V^R, E^R, f)$ is given we can construct the *robust configuration* $\mathcal{C}$ for the network coding with the state-of-the art coding strategies [28, 27]. Thus, in the rest of the dissertation I am only interested in the properties of the allocation of the optimal *survivable routing* $R = (V^R, E^R, f)$ of a connection to address the capacity efficiency issue of the dedicated protection approaches ensuring instantaneous recovery.

Thesis 1. *[C1, J1, C2] I proposed a new network coding based method called SRNC (Survivable Routing with Network Coding) which ensures* robust instantaneous recovery *for the arbitrary failure scenario. I proved that the problem is NP-complete at any finite user data splitting. According to that I proposed an Integer Linear Program (ILP) to solve the problem. For the single link failures scenario I proposed a capacity efficient, simple and low-complexity diversity coding based survivable routing method providing robust instantaneous recovery called SRDC (Survivable Routing with Diversity Coding). I investigated the method under capacity and node capability constraints (Table 2). I proposed several ILPs and heuristics to solve the different subproblems. I proved that the* $1+1$ *dedicated protection is a* $4/3$ *approximation of the subproblems without the capacity constraints (ICAN and ICCN). I showed through an example that the* $1+1$ *does not approximate the problem with capacity constraints (CCAN and CCCN). For the ICAN subproblem I proposed an approximation algorithm.*

The problems input is the following. Given is the input $\mathcal{I} = \{G, C, \mathcal{F}\}$ instance of the SRNC problem, where each connection is routed promptly after the request arrives to the edge router i.e., dynamic routing is assumed. In that time instant the actual topology of the network is denoted by $G = (V, E)$. On each edge $e \in E$ a non-negative cost function ($c : E \to R^+$) is defined, the free capacity ($k : E \to R^+$) and the delay value ($d : E \to R^+$) is given. A traffic demand $C = (s, t, b, D)$ consists of the source node s and sink node t, the bandwidth requirement ($b \in \mathbb{N}$) and the delay bound (D). Finally, the set of failure patterns, denoted as $\mathcal{F}$, is given. For each failure pattern, a failure subgraph is created: $\forall f \in \mathcal{F} : G_f = (V, E_f)$, where E_f is obtained by removing the edges in f from E. We assume that each failure pattern in $\mathcal{F}$ is *protectable*, i.e., each G_f is $s - d$ connected (the notations are summarized in Table 1).

It is important to emphasis that creating these auxiliary graphs is required only for a arbitrary failure list scenario, while in case of single link failure scenario other auxiliary graphs are utilized.

Note that while the arbitrary failure scenario may require network coding inside the network (i.e., node capabilities Figure 1 (a)-(d) are required) by the single links failure scenario network coding has to be performed only in the source and destination nodes i.e., the node capabilities in Figure 1 (a)-(c) are enough [28, 27].

4.1.1 Survivability of Transport Network for the Arbitrary Failure Scenario

Thesis 1.1 (SRNC complexity). *[J1, C2] I proved that by an arbitrary failure scenario the SRNC problem is NP-complete at any finite user data splitting. According to that I proposed an ILP to solve the problem.*

In particular I showed that finding an optimal SRNC solution in terms of minimizing the bandwidth usage assuming that the flow can be divided into two (or more generally into a finite integer T) parts is NP-complete. In order to prove the NP-completeness of the SRNC I presented a polynomial time transformation of the NP-complete GDP-R to SRNC [C2]. The poof is based on a graph transformation. Previously the NP-completeness of GDP-R was proven with the polynomial time transformation of the Steiner Forest problem (which is proven to be NP-hard [8]).

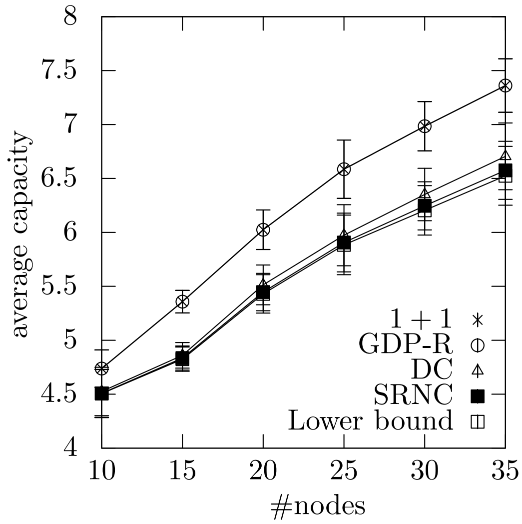
As the SRNC problem is NP-complete, I presented the corresponding ILP. In SRNC our goal is to obtain a solution $x \in \mathcal{X}_{\mathcal{I}}$ which minimizes the bandwidth cost

$$\min \sum_{\forall e \in E} c(e) \cdot \frac{b(e)}{b \cdot 2}. \qquad (2)$$

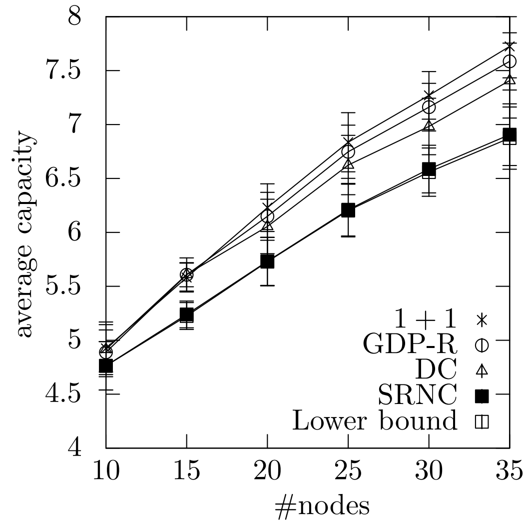
where $b(e)/2 \leq k(e)$ is the bandwidth reserved in the solution on link $e \in E$. The division by two is necessary since the user data is divided into two parts (A and B) and is routed individually through the network.

Table 1: Notation list for SRNC and SRDC

Notations	Description
$G = (V, E, k, c, d)$	the undirected graph representation V is the node and E is the edge set c is the cost function defined on $e \in E$ k is free capacity along link $e \in E$ d is the delay function defined on $e \in E$
$C = (s, t, b, D)$	source, destination and requested bandwidth and the delay bound of the dynamically arrived traffic demand, respectively
$\mathcal{F}$	failure patterns the connection needs to be resilient against
$G_f = (V, E_f)$	failure graph obtained by removing the failed edges in $f \in \mathcal{F}$ from E: $E_f = E \setminus f$
$R = (V^R, E^R, f)$	the survivable routing for the connection where $V^R \subseteq V$ is the set of nodes , $E^R \subseteq E$ is the set of edges, and $\forall e \in E^R : f(e) \leq k(e)$ are the flow values
$\mathcal{C}$	the robust configuration
$G^* = (V, E^*, c)$	directed multi-graph with edge set E^*, where all edge in $G = (V, E, k, c)$ are replaced by $k(e)$ parallel edges each with cost $c(e)$
$R^* = (V^{R^*}, E^{R^*})$	survivable routing of a connection in $G^* = (V, E^*, c)$, which is a DAG
A, B	the data parts after dividing the user flow
$A \oplus B$	the redundancy data part added with the XOR operation in s
$E_A, E_B, E_{A \oplus B}$	routing DAGs A, B and for $A \oplus B$



(a) Single link failures (degree 3.2)



(b) Multiple link failures (degree 3.2)

Figure 2: Average capacity reserved for a connection for different $\mathcal{F}$ failure scenarios and in networks $G = (V, E)$ with varying sizes [J1].

The program iterates through all the $\forall f \in \mathcal{F} : G_f = (V, E_f)$ auxiliary graphs, taking into account all the arbitrary failure patterns and setting up the proper variables i.e., creating the final solution. The main notations used are summarized in Table 1.

It has to be emphasized that I showed through extensive simulations that when dividing the user data into two parts, the theoretical lower bound (*Lower bound)* i.e., the GDP with network coding and arbitrary flow division can be very well approximated (see Figure 2).

4.1.2 Survivability of Transport Network for the Single Link Failure Scenario

Single link failures are the most common failures in transport networks, more than 70% [13] of all failures affecting links in the backbone networks are single link failures. Besides the $1+1$ dedicated protection the so called Diversity Coding (DC) [26] is the best candidate to protect single link failures. By DC redundancy is added at the source by sending a coded data packet ($A \oplus B$) on a third disjoint path constructed from the data of the two working paths (A and B, respectively) belonging to the

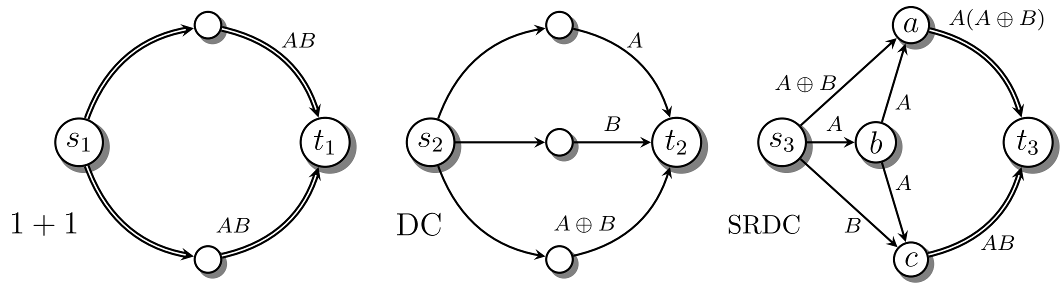


Figure 3: The illustration of the structure of the different protection approaches providing robust and instantaneous recovery. Double edges forward the whole user data, single edges the half of it [J3].

same connection. The DC improves the resource consumption of $1+1$ while providing instantaneous recovery in a simple manner [7]. However, the DC has high connectivity requirements, i.e., requires an at least 3-connected network, which is rarely the case in today's transport networks. An illustration of the different structures is presented in Figure 3.

Recently, the papers [28] and [27] made important steps to remedy the connectivity problem of diversity coding, and enable it to be an alternative for $1 + 1$ in transport network protection. It was shown that a minimum cost survivable routing with diversity coding can be decomposed into three Directed Acyclic Graphs (DAG), *which preserves one of the most important features of* $1+1$, *i.e., simplicity.* However, the issue of finding a minimum cost survivable routing was not tackled and addressed in these works.

This last step towards the practical implementation is discussed in the first part of my dissertation, namely how to fill the gap and solve the survivable routing problem in order to create a capacity efficient, simple and low-complexity diversity coding based survivable routing method providing robust and instantaneous recovery, which could replace the $1 + 1$.

The SRDC problem itself which is a low-complexity diversity coding based survivable routing method providing robust instantaneous recovery can be divided into four subproblems according to the technological constraints in the network. I investigated the SRDC under capacity and node capability constraints.

The problem formulations and their practical relevance is the following:

- ICAN (Infinite Capacity and All Node capabilities): In this case there are neither capacity nor node capability constraints in the network. “Infinite” capacity means that each link has the free capacity corresponding to the demand itself i.e., $k(e) \geq b$. In other words, there are no bottleneck links in the network, which can not accommodate the entire demand. Furthermore, we assume that each node is capable to perform the splitting and merging action. In this network scenario the network is not overloaded (i.e., no bottleneck links) and all the nodes received the necessary hardware or software updates i.e., all nodes are able to perform the splitting and merging actions. Note that in SDN networks a software update is enough to achieve the splitter and merger capabilities.

- ICCN (Infinite Capacity and Constrained Node capabilities): In this scenario there are no capacity constraints, but not all nodes are capable to perform the splitting and merging action i.e., the node capabilities are constrained/restricted. This network scenario belongs to the case where the network is not overloaded (i.e., no bottleneck links) but not all the nodes received the necessary hardware or software updates i.e., not all nodes are able to perform the splitting and merging actions. For example this scenario can occur if the operator decides to prefer incremental deployment.

- CCAN (Constrained Capacity and All Node capabilities): In this case all nodes are capable to perform the splitting and merging action, but there are bottleneck links in the networks i.e., there are capacity constraints. In this case the operator already updated all the nodes, however the network is overloaded i.e., some links have very limited free capacity.

- CCCN (Constrained Capacity and Constrained Node capabilities): In this scenario there are both capacity and node capability constraints in the network. In other words, there are bottleneck links in the network and not all the nodes received the necessary hardware or software updates. This can be considered as the worst case scenario.

It is already proven that if there are no capacity nor node constraints in the network (i.e., ICAN) we can find a solution in polynomial time with the help of SRDC-IA [C1]. The complexity of the capacity constraint or node constraint case i.e., CCAN or ICCN is still an open issue. However, it has been proven that if there are capacity and node constraints too i.e., CCCN (Constrained Capacity and Constrained Node capabilities) then the problem is NP-complete [C1]. These results are the existing results and are highlighted with the gray background in Table 2. All other results are the novel contributions of this dissertation (denoted with the white background) and will be discussed in details in the remainder of the chapter. An overview of the problems and results is given in Table 2.

Note that the polynomial time algorithm presented for the ICAN i.e., SRDC-IA can not be utilized for the other subporblems. For the ICCN problem I present an approximation algorithm the SRDC-IC. The algorithm is similar to the SRDC-IA, however virtual edges - which represent the minimum cost disjoint path-pair - are added only between the possible splitter-merger pairs. Furthermore note that the SRDC-IC does not deliver the optimal solution. An example is presented in Figure 4, where only node m is upgraded i.e., only node m can be a splitter or merger beside the source (s) and destination (t) node. If Diversity Coding (DC) is used, the total cost of the solution is 22, since the user data is sent along three edge disjoint paths (i.e., the sparsely dotted (5), the dashed (5) and $s \to v_{10} \to v_{11} \to t$ paths). If the SRDC-IC is used the cost of the solution is 20 (twice the cost of the sparsely dotted (5) and the dashed (5) paths), the same as by $1+1$. However, the optimal solution is 19 (dotted path (5), dashed path (5) and densely dotted path (9)), as shown in the Figure 4. Note that between nodes v_4 and m two units of data are transferred to get to the merging point in the network. We see that the SRDC-IC solution is not

Table 2: SRDC overview. The set of splitters and mergers is denoted as $\mathcal{P}$ and $\mathcal{M}$, respectively. The results with the gray background are existing approaches in the literature [C1], the novel contributions of this dissertation are denoted with the white background.

	No capacity constraints	Capacity constraints
$\mathcal{P} = V$, $\mathcal{M} = V$	*ICAN* SRDC-IA Polynomial time solvable	*CCAN* ILP + SRDC-CA -
$\mathcal{P} \subset V$, $\mathcal{M} \subset V$	*ICCN* ILP + SRDC-IC (Approx. algorithm) -	*CCCN* ILP + SRDC-CC NP-complete

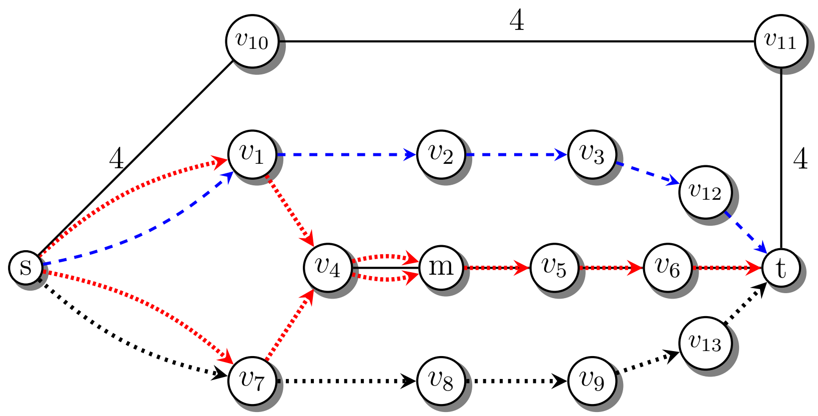


Figure 4: The optimal survivable routing solution of ICCN (total cost: 19). Beside the source and destination nodes, m is a potential merger/splitter node. If not displayed otherwise the edge costs are unit.

optimal, however, it approximates the problem, in the same way the $1+1$ does.

Thesis 1.2 (**SRDC Integer Linear Programs**). *[C1] I proposed a fast running ILP in order to solve the CCAN problem. I extended this ILP to solve the node capability constrained problem i.e., the CCCN problem, too.*

Of course, the running time of the CCCN ILP is considerably longer, since nodal constraints are also to be considered.

Thesis 1.3 (**SRDC heuristics**). *[C1] I proposed two heuristics for the CCCN and CCAN problems, i.e., the SRDC-CA which always provides a feasible solution (if exists) and the SRDC-CC which provides a faster and more efficient solution for most practical problem instances.*

This means the SRDC-CA heuristics always provide a feasible solution (if a solution exists) for the CCAN subproblem, however, if there are also node capability constraints in the network (i.e., CCCN subproblem), then this heuristic is no longer applicable. In this case the SRDC-CC should be used. The SRDC-CC does not guarantee a solution (since even deciding if there exists a solution for CCCN is already NP-complete as shown in [C1]), however it is extremely capacity efficient (the capacity consumption is near to the optimum i.e., *Lower Bound*, see Figure 5) and has a low blocking probability. For both heuristics, the [Theorem 2] presented in [28] is utilized, which states that finding a three-value flow in a so-called reduced capacity graph is equivalent with finding a single link failure survivable routing.

The main idea of the algorithm is to find a minimum cost $s-t$ flow with value of 3 in a special reduced capacity multigraph $G_{rs} = (V, E', \bar{c}, c_s)$. The auxiliary graph $G_{rs} = (V, E', \bar{c}_s, c_s)$ is created using the reduced capacities: $\bar{c}(e) = \min\{1, \bar{c}(e)\}$. In the transformation, the $e \in E$ edges of G are added to E' with their original cost $c(e)$. Based on its capacity $k(e)$ of $e \in E$, we add a parallel edge $e_n = (u, v)$ (called *extra edge*) to E' if $2 \leq k(e)$ with reduced capacity $\bar{c}(e_n) = 0.5$ and with cost $c_s(e_n) = c(e) \cdot \alpha$. This means that "creating islands" (i.e., using the extra edge e_n with $k'(e_n) = 0.5$ capacity) is penalized via a scaling factor (i.e., the extra edges have a higher cost $c_s(e_n) = c(e) \cdot \alpha, \alpha \geq 1$).

Of course choosing a proper α for a network is essential. For this purpose, intuitively we used the ratio $\alpha = \frac{|E|}{|V|}$, which is corresponding to the density (to be specific, it is half of the average nodal degree) of the network. In a denser network it is more likely to have 3 edge-disjoint paths (DC like solution) with a relatively short third path, which might be the optimal routing DAGs. Therefore, the cost of the extra edge is relatively high (α is high) in order to avoid creation of islands. Meanwhile in sparse networks, it is more likely that the third path (if it exists at all) is really long, and not beneficial to use ($1+1$ like solution). Of course sometimes the solution is identical with the DC or $1+1$ solution (i.e., the two extremes), depending on which is more cost (resource) efficient. But the truly valuable, new, and capacity-efficient solutions are those where we can extend the properties of DC to 2-connected graphs. This means that the solution of SRDC-CA are the three routing DAGs and not just two or three disjoint paths as by $1+1$ and DC. Furthermore the SRDC-CA heuristics is scalable and fast, so it can be very beneficial for large networks to use. The effectiveness of heuristics was verified by simulations [C1] (see Figure 5). As I have already mentioned, if the splitting and merging nodes can not be arbitrary, the problem gets even more difficult. The complexity of the ICCN problem is still an open question, but the CCCN is proved to be NP-complete [C1]. Nonetheless, the SRDC-CC heuristic is also based on a reduced capacity graph defined in [28]. However, we do not look for three-valued flow in the network, but for three paths, in a 2 step approach. The SRDC-CC heuristics is also scalable and even more fast running i.e., for large networks the use of SRDC-CC is very beneficial. The efficiency of heuristics was verified through simulations, where the SRDC-CC showed excellent results (see Figure 5).

Besides the survivable routing itself, I investigated the approximability of the problem, too. We know that when dividing the user data into two parts the $1+1$ is a 2-approximation of the survivable routing problem [6]. I proved that for the SRDC problem without capacity constraints the $1+1$ is a $4/3$ approximation.

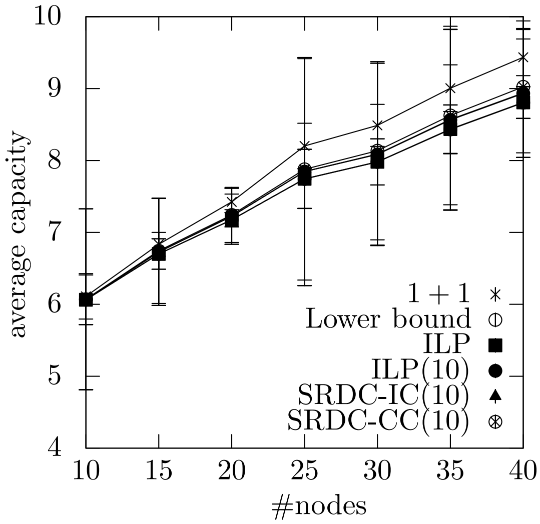
Thesis 1.4 (SRDC approximation). *I proved that the $1+1$ dedicated protection is a 4/3 approximation of the subproblems without the capacity constraints (ICAN and ICCN). I showed through an example that the $1+1$ does not approximate the problem with capacity constraints (CCAN and CCCN). For the ICCN subproblem I proposed an $4/3$ approximation algorithm.*

The proof is based on the SRDC-ICAN polynomial time algorithm that can be traced back to finding three paths in an auxiliary graph $\widehat{G}$. Using the inequalities between $1+1$ (2 paths) and SRDC-IA (ICAN) (3 paths), we prove algebraically that $1+1$ is a $4/3$ approximation for the SRDC with infinite capacity problems.

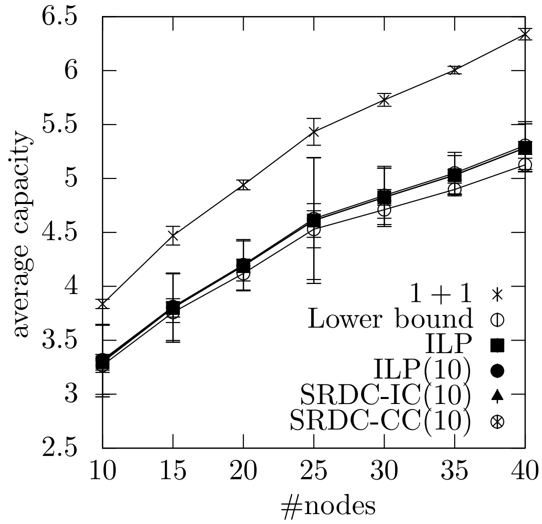
In order to show that in a capacity constraint case, the $1+1$ does not approximate the SRDC problem, we present a graph construction (Figure 6). Note that in Figure 6 only the double edges have the capacity of 2. In other words, these edges can accommodate the entire demand ($D = (s, t, 2)$) i.e., only these edges can be used by $1+1$. Furthermore, the single edges have a free capacity of 1 and can be only used by SRDC. This results in graph where the cost of $1+1$ is dependent on n (number of edges) i.e., the cost is $4+n$, and the cost of SRDC is always 6 (is independent from n), proving that the $1+1$ does not approximate the SRDC in a capacity constraint case.

4.2 Delay Aware Survivable Routing for the Single Link Failure Scenario

With the proliferation of multi-media and streaming applications, the end-to-end delay characteristics of the connections are getting more and more into the spotlight. The new applications (e.g., telesurgery, stock market, VoIP, etc.) are not only highly



(a) Optimality gap in the node constraint case in sparse graphs



(b) Optimality gap in the node constraint case in maxplan graphs

Figure 5: Bandwidth cost in sparse (average nodal degree between 2.4 and 3.2) and maximal planar (maxplan) graphs (average nodal degree between 4.2 and 5.7) in the infinite capacity, node capability constraint case.

delay sensitive, but they also require high resilience from the underlying network. Satisfying both constraints at the same time in a bandwidth-efficient way is unquestionably one of the most challenging tasks of service providers in transport networks.

To solve this problem, I propose new methods with low resource requirements that ensure *instantaneous recovery in a robust manner* while being able to satisfy *delay constraint even after a failure* occurs. To achieve this I incorporated delay constraints into the SRDC ICAN problem. This new problem is the so called Delay Aware Routing with Coding, i.e., DARC.

The DARC method is similar to SRDC and *DC* since the user data is sent over multiple paths (or DAGs) from the source to the target node. Therefore, the first step was to investigate the multi-path routing related delay problems. This problem has been thoroughly investigated in several network layers since multi-path routing has a real potential to provide the proper Quality-of-Service (QoS) level. With the help of the different disjoint paths we can achieve higher throughput, easier traffic control and high reliability.

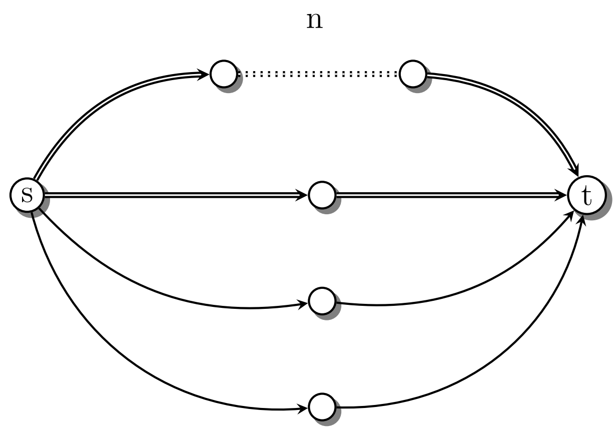


Figure 6: The counterexample of the approximability of SRDC-CCAN. The 1 + 1 is only capable to use the double edges, while SRDC can use all edges. From this follows that if we increase the number of edges n the difference in cost between the two solutions can be increased arbitrarily.

The two main multi-path routing related delay issues are the following [15]:

- The so called QoS (Quality of Service) routing: In this case our goal is to keep the sum of the delays of the paths under a given delay bound while minimizing the cost. The QoS routing problem is NP-complete even for a single path [8] and it is related to maintaining a certain QoS level, i.e, for example if we want to keep the end-to-end delay of the working path under a certain delay bound.

- Differential Delay Aware Routing (DD): In this case given is a delay bound for the difference between the path delays, while the cost function is minimized. The DD problem is also NP-complete, even for a single path i.e., for finding a path with overall delay between a given minimum and maximum bound [29]. The problem itself boils down to the issue of buffer size at the destination node for differential delay compensation.

Since the delay was not defined for DAGs, this was the first obstacle to overcome. After that I was able to introduce and solve the different delay related problems. When defining the delay of the DAGs I utilized the fact that the SRDC has a very strict structure i.e., that each routing DAG consists of series of paths and disjoint path-pairs, called **islands** ($\mathcal{I}$). Furthermore, each island is part of at most one routing

Table 3: The end-to-end delay difference on the two fastest routing DAGs upon a single link failure (wlog $\delta_{E_A} \leq \delta_{E_B} \leq \delta_{E_{A\oplus B}}$) [J2].

disrupted \ inc. delay	$\emptyset$	E_A	E_B	$E_{A\oplus B}$
$\emptyset$	$\lvert\delta_{E_A} - \delta_{E_B}\rvert$	$\lvert\delta_{E_B} - \min\{\Delta_{E_A}, \delta_{E_{A\oplus B}}\}\rvert$	$\lvert\delta_{E_A} - \min\{\Delta_{E_B}, \delta_{E_{A\oplus B}}\}\rvert$	$\lvert\delta_{E_A} - \delta_{E_B}\rvert$
E_A	$\lvert\delta_{E_B} - \delta_{E_{A\oplus B}}\rvert$	–	$\lvert\Delta_{E_B} - \delta_{E_{A\oplus B}}\rvert$	$\lvert\delta_{E_B} - \Delta_{E_{A\oplus B}}\rvert$
E_B	$\lvert\delta_{E_A} - \delta_{E_{A\oplus B}}\rvert$	$\lvert\Delta_{E_A} - \delta_{E_{A\oplus B}}\rvert$	–	$\lvert\delta_{E_A} - \Delta_{E_{A\oplus B}}\rvert$
$E_{A\oplus B}$	$\lvert\delta_{E_A} - \delta_{E_B}\rvert$	$\lvert\Delta_{E_A} - \delta_{E_B}\rvert$	$\lvert\delta_{E_A} - \Delta_{E_B}\rvert$	–

DAG (see SRDC). It is important to emphasis that also by *DARC* the network coding can be done in the source and destination nodes. In other words we do not have to upgrade the network node capabilities in order to provide delay aware robust instantaneous recovery (in a simple manner).

Thesis 2. *[C3, J2] I defined the delay of routing DAGs. I introduced the DARC-QoS and DARC-DD problems. I proved that both problems, namely the DARC-QoS and the DARC-DD, are NP-complete. Furthermore, I have proved that the DARC-DD problem can not be approximated within n^ϵ for any $\epsilon < 1$ in Hamilton graphs. I presented ILPs to solve both problems, and introduced two effective heuristics to solve DARC-DD.*

In order to capture the before- and after-failure delay characteristics of the routing DAGs, I introduced two delay values for each island I: $d^I_{min} = \sum_{e \in I_{min}} d(e)$ corresponding to the delay of the island in the failure-less state (i.e., the lower delay path); and the delay difference between the two disjoint paths $\Delta^I = \sum_{e \in I_{max}} d(e) - \sum_{e \in I_{min}} d(e)$ corresponding to the delay increase between the splitter and merger node of the island upon a failure occurs on the lower delay path. Thus, the *end-to-end delay of a routing DAG* can be modeled as

$$\delta_{E_j} = \sum_{P \in \mathcal{P}_{E_j}} \sum_{e \in P} d(e) + \sum_{I \in \mathcal{I}_{E_j}} d^I_{min} \tag{3}$$

in the failure-less state, while it increases to

$$\Delta_{E_j} = \delta_{E_j} + \max_{I \in \mathcal{I}_{E_j}} \Delta^I \tag{4}$$

in worst case upon a failure along the island with the largest delay difference between its two paths[1].

This means in a failure-less state the delay of a routing DAG E_j corresponds to the sum of minimum delays on the paths and islands. However in a case of a failure we assume the worst case i.e., that the island with the largest delay difference between its two paths fails. It is important that both the ILPs and heuristics run on a auxiliary graph in which the islands are represented by virtual edges. To create such a graph we improve the graph transformation presented in [C1]. The auxiliary (multi-)graph $G^* = (V, E^*, c^*, d_{min}, \Delta)$ is created, so that the link set E^* contains the *original links* of G. Additional *virtual links* $e_{(u,v)}$ are added between every distinct node-pairs for which a link-disjoint path-pair exist. The cost of these virtual edges $c^*(e_{(u,v)})$ is set to the cost of a minimum cost link-disjoint path-pair between nodes u and v in G. In addition to the previous transformation for SRDC, in DARC we have to capture the routing DAG delays in Eq (3)-(4). Thus, two variables d^I_{min} and Δ^I are introduced for each virtual link (island) $I = e_{(u,v)}$ (i.e., the delay of I_{min} and the delay difference between I_{min} and I_{max}). For original links $e \in E$ we define $d^e_{min} = d(e)$, and $\Delta^e = 0$. In such a graph the QoS and DD problems can be traced back to finding three link-disjoint $s - t$ paths with minimum total cost. We note here that, we can use the transformed graph to run algorithms, but owing to the correlation between the links and link parameters *the hardness results can not be directly transformed back to our original survivable routing problem.* Thus, further investigation was required to show the complexity of survivable QoS routing and DD aware routing problems.

Thesis 2.1 (**QoS complexity**). *[C3, J2] I proved that the DARC-QoS problem is NP-complete.*

The proof is based on the reduction from the Three-Way Partition problem [8], which is known to be NP-hard.

Thesis 2.2 (**DD complexity**). *[C3, J2] I proved that the DARC-DD problem is NP-complete.*

The proof is based on the polynomial reduction (graph construction) from the Longest Path Problem [8] , which is known to be NP-complete.

[1] Note that, $\Delta_{E_j} - \delta_{E_j}$ gives also the largest jitter we might expect along the routing DAG caused by protection switching.

An important difference between the QoS and DD problems is, that by DD we have to satisfy very strict differential delay constraints, hence the cost of the results can become secondary. This means if the ILP is not formulated correctly loops or node-disjoint cycles can be built just to satisfy the delay constraints. Of course these solutions are not acceptable, even if strictly they satisfy the constraints. To avoid this, I used the so called voltage analysis method in the ILP and for the loops I added new flow constraints. In addition note that the DARC-DD problem aims two keep the differential delay difference between the to fastest DAGs, hence we have to satisfy all the constraints given in Table 3.

Thesis 2.3 (**Integer Linear Programs**). *[C3, J2] Based on the novel definition of routing DAG delays I presented ILPs to solve the QoS and DD problem.*

Thesis 2.4 (**Heuristics**). *[J2] Because of the slow running time of the ILP, I suggested two fast heuristics (cost and delay based) by modifying the SPLIT [11] approach to be able to solve the DD problem.*

Since the problem is hard to approximate, it is legitimate to use a fast heuristic approach, which leverages the computational complexity of the presented ILP. The first step of the algorithm is to find (first $k = 3$) k-shortest link-disjoint paths. As a second step, the algorithm identifies the intermediate node on which the most paths go through, denote this node with v_x. Then, the paths that pass through node v_x are split into segments ($s - v_x$ and $v_x - t$ where s a source and t the target node). These segments are glued together in such a way that the least delayed $v_x - t$ segment is assigned to the largest delay $s - v_x$ segment. The edges used are deleted, and then we return to the second step until the graph is not empty. We create triplets from the set of paths and check if they satisfy the delay constraints. If yes, then we save the triplet to the possible solutions set. The value of k is incremented iteratively, as long as we can find k link-disjoint paths. At the end of the algorithm, we select from the solution set (i.e., from the triplet) the one with the lowest possible cost. Note that this algorithm is running on the auxiliary graph.

Thesis 2.5 (**DD approximability**). *[C3, J2] I have proved that the DARC-DD problem can not be approximated within n^ϵ for any $\epsilon < 1$ in Hamilton graphs.*

The approximability question boils down to the question of finding path longer than $n - n^\epsilon$ in Hamiltonian graphs, which is already proven to be NP-hard. This method was used by Srivastava in [33][Section IV C] to prove that the DDR (Differential Delay Routing) problem can not be approximated within a factor of n^ϵ for any $\epsilon < 1$ for Hamiltonian graphs. Despite the relevant differences of DDR and DARC-DD, the proofs follow similar reasoning.

4.3 Switching Action Based Failure Localization in Transport Networks

In the first part of my dissertation I proposed new methods with low resource requirements that ensure *instantaneous recovery in a robust manner* while being able to satisfy *delay constraints even after a failure* occurs. In other words, the first part of my dissertation was all about the survivable routing under different technological constraints i.e., capacity, node capability and delay constraints. This means we focused on protection and not on failures localization, we reserved and provisioned the necessary backup resources in advance in order to protect the connection in case of a certain failure pattern (e.g. single link failures or arbitrary SRLG lists).

In order to reduce the resource consumption even further, we have to turn to the shared protection approaches. Until now shared protection approaches were considered to be resource efficient, however, they had a long recovery time due to long failure localization and notification times. That is why reducing the recovery has been a subject of investigation since long.

In order to enable fast *failure localization* in all-optical networks, supervisory lightpaths (S-LP, monitoring trails or m-trails) have been introduced [5]. An m-trail consists of a pair of lightpaths along a common physical route (in opposite directions), and is purely used for monitoring the on/off status of the physical links along the route at the end-nodes.

First centralized all-optical m-trail allocation schemes have been introduced [4, 30, 31, 21], where a central failure manager is responsible for unambiguous failure localization based on the alarm messages of the (small set of) disrupted m-trails. Later, the concept of Local Unambiguous Failure Localization (L-UFL) was introduced [4], where the goal was to completely eliminate signaling from the failure localization phase with a properly allocated set of m-trails at a distinguished node of the network where these S-LPs terminate, called *L-UFL capable* node. However, the L-UFL node (or the failure manager) is still needed to notify the switching nodes about the failure with electrical layer messages.

The goal in Network-Wide Local Unambiguous Failure Localization (NWL-UFL) [30] was to enable all-optical signaling-free failure restoration (i.e., eliminate messages in both failure localization and failure notification phases) by making all nodes in the network L-UFL capable under any single link failure. After the failure is localized at each node of the network, the corresponding switching actions can be performed immediately. However, the lengthy m-trails (spanning trees [30] to be specific) cause implementation issues and higher monitoring latencies. In addition, we unambiguously identify all the link failures, although not all link failures have to be unambiguously localized at each node, e.g., if that particular node does not need to perform protection switching action for a failure.

In order to improve these drawbacks of NWL-UFL, Global Neighborhood Failure Localization (G-NFL) has been proposed in [31, 21]. The main idea is to unambiguously localize only a small set of link failures that a particular node needs to respond to in the restoration process via short m-trails. These identifiable links are referred to as the *neighborhood* of the node. Hence, G-NFL is currently the most efficient signaling-free unambiguous link failure localization framework, as it allows the *shortest possible W-LPs and P-LPs to be used* in the data plane, as well as short S-LPs (simple or non-simple paths) to be allocated in the control plane. Thus, upon request an S-LP can be turned into a W-LP, while the status of W-LPs can be used for failure monitoring purposes as well. The goal of previous m-trail allocation methods was unambiguous localization of single link failures.

In my dissertation I introduce a *novel framework which relaxes the requirement of unambiguous link failure localization*, focusing not on link failures but rather on the identifiable *protection switching actions*. This means for example, if the same protection switching action belongs to two different link failures, then we do not have to localize the two failures unambiguously (separately), we only need to identify that some link belonging to a given switching group has failed, in order to initiate the proper protection switch. This new framework is the so called Advanced Global

Neighborhood Failure Localization (AG-NFL). With this new method we are able to reduce the resource consumption significantly.

An example is presented in Fig. 7(a), where W-LP W_1 and two P-LPs $P^{(v_1 \to v_3)}$ and $P^{(v_3 \to v_4)}$ are given. If (v_1, v_2) or (v_2, v_3) fails then $P^{(v_1 \to v_3)}$, and if (v_3, v_4) fails then $P^{(v_3 \to v_4)}$ should be activated, respectively. In this case the AG-NFL needs only two m-trails – T_1 and T_2 in Fig. 7(b) – to enable all nodes included in the recovery process of W_1 to *perform the proper switching action* (e.g., if T_1 and W_1 are down then $P^{(v_1 \to v_3)}$ is activated at nodes v_1, v_3, v_5, and v_6). However, G-NFL would require further S-LPs (shown in Fig. 7(c)) to unambiguously identify neighborhood links $(v_1, v_2), (v_2, v_3)$ at v_6 (in total 5 S-LPs). We can observe that in node v_6 the AG-NFL does not distinguish between the link failures (v_1, v_2) and (v_2, v_3), since the same switching action has to be activated i.e., $P_1^{(v_1 \to v_3)}$. Also, the AG-NFL does not differentiate between the failure of (v_1, v_5) and (v_5, v_6) and the failure of (v_1, v_2) and (v_2, v_3). This is possible because the AG-NFL allows the so-called *"non-disruptive" switching*, that is, if (v_1, v_5) or (v_5, v_6) fails, then $P_1^{(v_1 \to v_3)}$ can be performed in node v_6, we switch only pre-allocated protection capacity, nothing will happen, just some switching energy is wasted. Of course, such an action is not allowed to disrupt any working path, that is why we have to be careful when creating the switching link groups. We see that the gain is more than 50% (2 S-LPs instead of 5).

This advantage of AG-NFL is based on the following two observations:

- **Observation (i):** If exactly the same switching action belongs to two links in the neighborhood at node v, then we do not have to distinguish their failures from each other to perform the proper switching action. In other words, these links are a *working segment* of the corresponding W-LP (e.g., links $(v_1, v_2), (v_2, v_3)$ at v_6 in Fig. 7(a)).

- **Observation (ii):** Link failures in the neighborhood of a node with only *"non-disruptive" switching actions* do not have to be distinguished from links outside of the neighborhood (i.e., links to which no switching action belongs). For example, links $(v_1, v_5), (v_5, v_6)$ do not need to be distinguished from neighborhood links $(v_1, v_2), (v_2, v_3)$ at v_6 in Fig. 7(a), as v_6 is an intermediate node of $P^{(v_1 \to v_3)}$. Thus, false switching on at v_6 will not disrupt W-LP W_1.

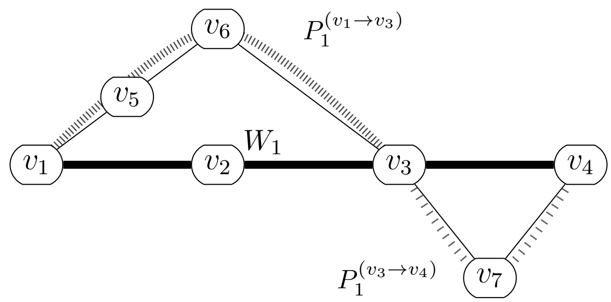
Table 4: Summary of the need of m-trail reconfiguration upon data-plane changes [J4] [31]

	SOD-IO	SOD-O	LOD-IO	LOD-O
W-LP setup	X	X		
W-LP teardown	X		X	

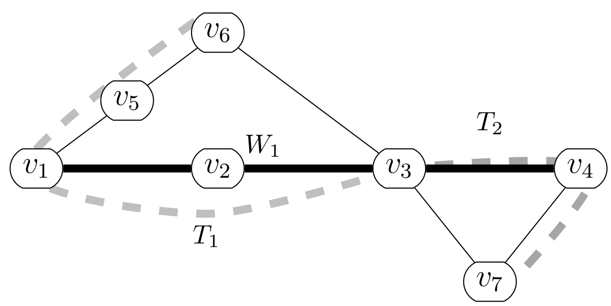
Based on these principles, the neighborhood is split into two sets, that is, the group of links that require exact and approximate identification. With this new, even finer classification, we can significantly reduce the resource requirements for the failure localization for any shared protection. But more importantly, this new all-optical local failure localization framework, enables fast restoration even for the shared protection approaches, reducing the failure recovery time significantly. Note that due to the nature of the shared protections optical cross-connect (OXC) reconfiguration is always needed. This means *robust recovery can not be achieved* i.e., the simplicity level of $1 + 1$ can not be maintained.

In addition I examined several m-trail design aspects and problems with different data plane dependencies [31]. The first aspect is, if the AG-NFL only utilizes the status information of the S-LPs for switching link group failure localization we call it out-of-band (-O) failure localization, if both the status information of W-LPs and S-LPs are considered we refer to it as in- and out-of-band (-IO) monitoring, as in-band status information of W-LPs carrying data plane traffic is also considered in the failure localization. On the other hand, depending on whether we are optimizing the current traffic matrix (Strictly On-Demand (SOD)) or all the traffic on the future (Loosely On-Demand (LOD)), we get another problem [31]. We summarized the above m-trail scenarios and their dependency on the data plane in Table 4.

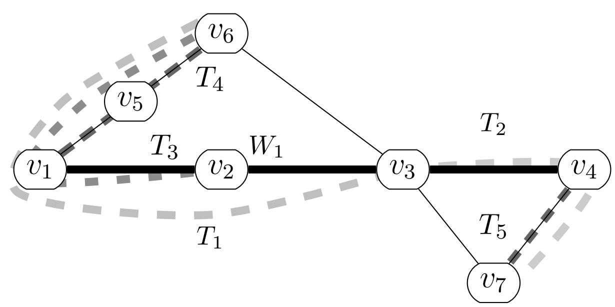
Thesis 3. *[C4, J4] I introduced a novel local failure localization framework called Advanced Global Neighborhood Failure Localization, which enables ultra fast recovery even for shared protection methods. I introduced a new concept called forbidden link-pairs which generalizes several m-trail problems (AG-NFL, G-NFL, NWL-UFL) and*



(a) W-LP W_1 and the corresponding P-LPs.



(b) S-LPs T_1 and T_2.



(c) G-NFL solution (needs three extra S-LPs).

Figure 7: Monitoring the status of W_1 and S-LPs T_1, T_2 provides enough information for each node to perform proper protection switching action, while G-NFL needs additional m-trails $T_3 - T_5$ for unambiguous link failure localization. Intermediate nodes are able to monitor the status of the traversing m-trails and W-LPs.

reduces their solutions complexity. I proved that both the AG-NFL SOD-IO and LOD-IO are NP-complete. These are the first complexity results for single link failure localization with m-trails. To solve this problem, I proposed a new heuristic.

Thesis 3.1 (Generalized concept). *[J4] I introduced a new concept called forbidden link-pairs which generalizes several m-trail problems and reduces their solutions complexity. Based on this concept I proposed an efficient heuristic to solve these problems.*

Thesis 3.2 (AG-NFL complexity). *[J4] I proved that both the AG-NFL SOD-IO and LOD-IO are NP-complete.*

The NP-completeness poof is based on the Hamiltonian $s-t$ Path Problem [8, GT39] i.e., I proved that the AG-NFL is NP-complete with a help of a graph construction. To solve the problem I presented an efficient and fast Dijkstra based heuristic.

We can observe in the simulation results presented in Table 5 that the new framework reduces the resource consumption significantly, sometimes even 50%. Of course the gain depends on the m-trail design problem (SOD or LOD, -IO or -O).

5 Summary

In the first part of my dissertation I proposed new network coding based methods, which are able to satisfy the new complex QoS requirements i.e., I proposed methods with low resource requirements that ensure *instantaneous recovery in a robust manner* while being able to satisfy *delay constraint even after a failure* occurs. I compare the new methods with $1+1$ and other state-of-the-art protection approaches. These methods proved their capabilities in the real word networks i.e., **the methods were implemented in a European size SDN network, where they verified their effectiveness**.

In the second part of the dissertation I focus on the failure localization. I propose a new all-optical local failure localization framework, that enables fast restoration even for the shared protection approaches, reducing the failure recovery time and the resource consumption of the failure localization method significantly.

Table 5: Results on real network topologies from the Internet Topology Zoo [38]. In the SOD scenario 30% of $s-t$ pairs loaded. In the LOD scenario the same 30% of W-LPs are active, while forbidden link-pairs calculated for 100% traffic. AG-NFL columns contain worst case results with only A-N code collisions allowed (Obs. (ii)) in addition to G-NFL [J4].

Network topology			$\frac{\lVert T\rVert}{\lvert E\rvert}$								b							
			SOD-IO		SOD-O		LOD-IO		LOD-O		SOD-IO		SOD-O		LOD-IO		LOD-O	
Name	$\lvert V\rvert$	$\lvert E\rvert$	G-NFL	AG-NFL	G-NFL	AG-NFL	G-NFL	AG-NFL	G-NFL	AG-NFL	G-NFL	AG-NFL	G-NFL	AG-NFL	G-NFL	AG-NFL	G-NFL	AG-NFL
Abilane	11	14	2.14	1.14	3.71	2.92	2.14	1.57	3.71	3.00	13	8	26	23	13	10	26	22
Germany	17	25	2.20	1.32	3.48	2.60	2.52	1.48	3.96	3.12	26	19	45	39	29	19	49	43
BtEurope	17	30	1.50	0.86	2.90	2.20	2.03	1.63	3.36	2.90	25	13	52	39	31	24	56	48
AS6461	17	37	2.16	0.97	3.56	2.37	2.59	1.75	3.97	3.02	44	19	75	50	53	34	83	63
InternetMCI	18	32	2.53	1.03	4.12	2.75	3.09	1.81	4.59	3.68	38	16	70	50	45	27	74	63
AS1755	18	33	1.60	0.81	2.72	2.15	2.54	1.90	3.78	3.24	30	17	54	41	42	32	68	59
ChinaTelc	20	44	1.56	0.77	3.68	3.20	5.15	2.56	6.93	4.75	37	18	86	73	105	52	145	101
AS3967	21	36	3.19	1.94	5.02	4.02	3.41	2.16	5.38	4.52	52	33	89	71	55	36	94	79
BellSouth	21	36	0.75	0.36	2.27	2.02	0.75	0.55	2.27	2.11	16	7	54	48	16	10	54	48
AT&T	22	38	1.76	0.86	4.05	3.34	2.50	1.73	4.71	4.23	33	17	76	60	48	35	90	80
NSF	26	43	3.00	2.00	5.60	4.76	3.65	2.97	6.32	5.79	55	38	106	91	65	54	116	107
BICS	27	42	2.88	1.07	5.42	4.40	3.42	2.02	5.90	4.85	50	22	105	83	58	37	110	89
AS3257	27	64	2.34	1.39	4.46	3.50	2.68	1.85	4.79	3.98	78	44	150	115	88	56	157	125
AS1239	30	69	2.39	1.01	5.49	3.98	3.94	2.14	7.18	5.37	80	35	181	130	121	65	226	164
Arnes	31	47	2.17	1.12	5.08	3.93	2.53	1.70	5.29	4.48	47	24	112	86	52	36	112	95
Geant	31	49	2.65	1.69	5.32	4.20	3.18	2.28	5.73	5.02	56	39	120	97	64	48	125	111
Italy	33	56	1.69	0.96	3.96	3.37	1.76	1.17	4.00	3.71	47	30	116	99	47	35	116	105
BtNAmerica	36	76	3.21	2.01	6.93	5.48	4.03	2.69	7.63	6.22	109	68	239	184	134	87	257	204
BellCanada	39	55	3.25	2.12	7.76	5.80	3.56	2.45	8.16	6.20	62	45	156	121	68	53	163	129

References

[1] CPLEX. http://www.ilog.com/products/cplex/.

[2] GUROBI. Gurobi Optimization, Inc http://www.gurobi.com/products/gurobi-optimizer/gurobi-overview/.

[3] Lemon: A c++ library for efficient modeling and optimization in networks. Technical report, http://lemon.cs.elte.hu.

[4] S. Ahuja, S. Ramasubramanian, and M. Krunz. Single-link failure detection in all-optical networks using monitoring cycles and paths. *IEEE/ACM Transactions on Networking (TON)*, 17(4):1080–1093, 2009.

[5] P. Babarczi, J. Tapolcai, and P.-H. Ho. Adjacent link failure localization with monitoring trails in all-optical mesh networks. *IEEE/ACM Transactions on Networking*, 19(3):907–920, 2011.

[6] G. Brightwell, G. Oriolo, and F. B. Shepherd. Reserving resilient capacity in a network. *SIAM journal on discrete mathematics*, 14(4):524–539, 2001.

[7] A. Fumagalli and L. Valcarenghi. IP restoration vs. WDM protection: is there an optimal choice? *Network, IEEE*, 14(6):34–41, 2000.

[8] Garey, Michael R and Johnson, David S *Computers and intractability: a guide to NP-completeness*, WH Freeman New York, 1979.

[9] W. Grover, J. Doucette, M. Clouqueur, D. Leung, and D. Stamatelakis. New options and insights for survivable transport networks. *IEEE Commun. Mag.*, 40(1):34–41, Jan. 2002.

[10] P.-H. Ho, J. Tapolcai, and T. Cinkler. Segment shared protection in mesh communication networks with bandwidth guaranteed tunnels. *IEEE/ACM Transactions on Networking*, 12(6):1105–1118, December 2004.

[11] S. Huang, C. Martel, and B. Mukherjee. Survivable multipath provisioning with differential delay constraint in telecom mesh networks. *IEEE/ACM Transactions on Networking*, 19(3):644–656, 2011.

[12] L. Huawei Technologies Co. White paper on technological developments of optical networks. *white paper,Huawei Technologies Co., Ltd*, 2016.

[13] A. Markopoulou, G. Iannaccone, S. Bhattacharyya, C.-N. Chuah, and C. Diot. Characterization of failures in an ip backbone. In *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, volume 4, pages 2307–2317. IEEE, 2004.

[14] D. Papadimitriou and E. Mannie. Analysis of generalized multi-protocol label switching (gmpls)-based recovery mechanisms (including protection and restoration). Technical report, 2006.

[15] H. F. Salama, D. S. Reeves, and Y. Viniotis. A distributed algorithm for delay-constrained unicast routing. In *INFOCOM'97. Sixteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Driving the Information Revolution., Proceedings IEEE*, volume 1, pages 84–91. IEEE, 1997.

[16] J.P. Vasseur, M. Pickavet, and P. Demeester. *Network recovery: Protection and Restoration of Optical, SONET-SDH, IP, and MPLS.* Morgan Kaufmann Publishers, 2004.

[17] Musumeci, F. and Tornatore, M. and Pattavina, A. A Power Consumption Analysis for IP-Over-WDM Core Network Architectures. *Optical Society of America, Journal of Optical Communications and Networkin*, 4(2):108–117, 2012

[18] Y. Yano, T. Ono, K. Fukuchi, T. Ito, H. Yamazaki, M. Yamaguchi, and K. Emura. 2.6 Terabit/s WDM transmission experiment using optical duobinary coding. In *Optical Communication, 1996. ECOC'96. 22nd European Conference on*, volume 5, 1996.

[19] Konstantinos N. Georgakilas, Kostas Katrinis, Anna Tzanakaki, and Ole B. Madsen. Impact of Dual-Link Failures on Impairment-Aware Routed Networks. In *Transparent Optical Networks, 2010. Proceedings of 2010 12th International Conference on*. IEEE, 2010.

[20] A. Kotb and K. E. Zoiros. Performance analysis of all-optical xor gate with photonic crystal semiconductor optical amplifier-assisted mach–zehnder interferometer at 160 gb/s. *Optics Communications*, 402:511–517, 2017.

[21] J. Tapolcai, P.-H. Ho, P. Babarczi, and L. Rónyai. On achieving all-optical failure restoration via monitoring trails. In *Proc. 32nd IEEE INFOCOM*, pages 380–384, April 2013.

[22] J. Tapolcai, P.-H. Ho, L. Rónyai, P. Babarczi, and B. Wu. Failure localization for shared risk link groups in all-optical mesh networks using monitoring trails. *IEEE/OSA Journal of Lightwave Technology*, 29(10):1597–1606,

[23] J. Strand, AL Chiu, and R. Tkach. Issues for routing in the optical layer. *IEEE Communications Magazine*, 39(2):81–87, 2001.

[24] Ahlswede, R. and Cai, N. and Li, S.Y.R. and Yeung, R.W. Network information flow *IIEEE Transactions on Information Theory*, 46(4):1204–1216, 2000.

[25] Fragouli, C. and Le Boudec, J.Y. and Widmer, J. Network coding: an instant primer *ACM SIGCOMM Computer Communication Review*, 36(1):63–68, 2006.

[26] Ayanoglu, Ender and Chih-Lin, I and Gitlin, Richard D and Mazo, James E. Diversity coding for transparent self-healing and fault-tolerant communication networks *IEEE Transactions on Communications*, 41(11):1677–1686, 1993.

[27] Babarczi, Peter and Tapolcai, Janos and Rónyai, Lajos and Médard, Muriel Resilient Flow Decomposition of Unicast Connections with Network Coding *Proc. IEEE Intl. Symp. on Information Theory (ISIT)*,:116-120, 2014.

[28] Rouayheb, S. and Sprintson, A. and Georghiades, C. Robust Network Codes for Unicast Connections: A Case Study *IEEE/ACM Transactions on Networking*, 19(3):644–656, 2011.

[29] Ahuja, Satyajeet and Krunz, Marwan and Korkmaz, Turgay Optimal path selection for minimizing the differential delay in Ethernet-over-SONET, *Elsevier, Computer Networks*, 50(13):2349–2363, 2006.

[30] J. Tapolcai, P.-H. Ho, P. Babarczi, and L. Rónyai, “On signaling-free failure dependent restoration in all-optical mesh networks,” *IEEE/ACM Transactions on Networking*, vol. 22, no. 4, pp. 1067–1078, Aug 2014.

[31] J. Tapolcai, P. H. Ho, P. Babarczi, and L. Rónyai, “Neighborhood failure localization in all-optical networks via monitoring trails,” *IEEE/ACM Transactions on Networking*, vol. 23, no. 6, pp. 1719–1728, Dec 2015.

[32] A. Orda and A. Sprintson, “Efficient algorithms for computing disjoint QoS paths,” in *23th IEEE INFOCOM*, vol. 1, 2004.

[33] A. Srivastava, S. Acharya, M. Alicherry, B. Gupta, and P. Risbood, “Differential delay aware routing for Ethernet over SONET/SDH,” in *24th IEEE INFOCOM*, vol. 2, 2005, pp. 1117–1127.

[34] Georgios Ellinas, Eric Bouillet, Ramu Ramamurthy, Jean-Francois Labourdette, Sid Chaudhuri, and Krishna Bala. Routing and restoration architectures in mesh optical networks. *Optical Networks Magazine*, pages 91–106, January/February 2003.

[35] Tapolcai, János and Wu, Bin and Ho, Pin-Han and Rónyai, Lajos A novel approach for failure localization in all-optical mesh networks. *IEEE/ACM Transactions on Networking*, vol. 19, 2011 pages 275–285,

[36] Yadav, Ravinder and Aggarwal, Rinkle Rani Survey and Comparison of Optical Switch Fabrication Techniques and Architectures, arXiv preprint arXiv:1004.4481, 2010

[37] Ma, Xiaohua and Kuo, Geng-Sheng Optical switching technology comparison: optical MEMS vs. other technologies, *IEEE communications magazine*, vol. 41, 2003 pages S16–S23,

[38] S. Knight, H. X. Nguyen, N. Falkner, R. Bowden, and M. Roughan. The Internet Topology Zoo. http://www.topology-zoo.org.

References

Journal Publications

[J1] P. Babarczi, A. Pašić, J. Tapolcai, F. Németh, and B. Ladóczki. Instantaneous recovery of unicast connections in transport networks: Routing versus coding. *Elsevier Computer Networks*, 82(1):68–80, 2015. (6/4 = 1,5)

[J2] A. Pašić, P. Babarczi, and A. Kőrösi. Diversity coding-based survivable routing with QoS and differential delay bounds. *Optical Switching and Networking*, 23(2):118–128, 2017. (6/2 =3)

[J3] P. Babarczi, J. Tapolcai, A. Pašić, L. Rónyai, E. R. Bérczi-Kovács, and M. Médard. Diversity coding in two-connected networks. *IEEE/ACM Transactions on Networking*, 25(4):2308–2319, 2017. (6/5 =1.2)

[J4] A. Pašić, P. Babarczi, and J. Tapolcai. Unambiguous switching link group failure localization in all-optical networks. *Networks*, 70(4):327–341, 2017. (6/2 =3)

Conference Proceedings Publications

[C1] A. Pašić, J. Tapolcai, P. Babarczi, E. Bérczi-Kovács, Z. Király, and L. Rónyai. Survivable routing meets diversity coding. In *IFIP Networking*, pages 1–9, 2015. (3/5 = 0,6)

[C2] P. Babarczi, J. Tapolcai, A. Pašić, S. R. Darehchi, and P.-H. Ho. New addressing scheme to increase reliability in mpls with network coding. In *Design of Reliable Communication Networks (DRCN), 2013 9th International Conference on the*, pages 36–43. IEEE, 2013. (3/4 = 0,75)

[C3] A. Pašić and P. Babarczi. Delay aware survivable routing with network coding in software defined networks. In *Reliable Networks Design and Modeling (RNDM), 2015 7th International Workshop on*, pages 41–47. IEEE, 2015. (3/1 = 3)

[C4] A. Pašić and P. Babarczi. Switching link group failure localization via monitoring trails in all-optical networks. In *Resilient Networks Design and Modeling (RNDM), 2016 8th International Workshop on*, pages 92–99. IEEE, 2016. (3/1 = 3)

Other Publications

[C5] Tornatore, Massimo and André, Joao and Babarczi, Péter and Braun, Torsten and Følstad, Eirik and Heegaard, Poul and Hmaity, Ali and Furdek, Marija and Jorge, Luisa and Kmiecik, Wojciech and Mas Machucaand, Carmen and Martins, Lucia and Medeiros, Carmo and Musumeci, Francesco and Pašić, Alija and Rak, Jacek and Simpson, Steven and Travanca, Rui and Voyiatzis, Artemios. A survey on network resiliency methodologies against weather-based disruptions. In *Resilient Networks Design and Modeling (RNDM), 2016 8th International Workshop on*, pages 23–34. IEEE, 2016. (3/18 = 0.16)

[C6] Gomes, Teresa and Tapolcai, János and Esposito, Christian and Hutchison, David and Kuipers, Fernando and Rak, Jacek and de Sousa, Amaro and Iossifides, Athanasios and Travanca, Rui and André, Joao and Jorge, Lusa and Martins, Lcia and Ortiz Ugalde, Patricia and Pašić, Alija and Pezaros, Dimitrios and Jouet, Simon and Secci, Stefano and Tornatore, Massimo. A survey of strategies for communication networks to protect against large-scale natural disasters. In *Resilient Networks Design and Modeling (RNDM), 2016 8th International Workshop on*, pages 11–22. IEEE, 2016. (3/15 = 0.2)

[C7] Pašic, Alija and Girao-Silva, Rita and Vass, Balázs and Gomes, Teresa and Babarczi, Péter FRADIR: A Novel Framework for Disaster Resilience. In *Resilient Networks Design and Modeling (RNDM), 2018 10th International Workshop on*, pages 1–7. IEEE, 2018. (3/4 = 0.75)