Link Weight Optimization Models against Link Failure in Internet Protocol Networks

Kaptchouang Stephane
Department of Communication Engineering and Informatics
The University of Electro-Communications

A thesis submitted for the degree of

Doctor of Philosophy

March 2017
APPROVED BY SUPERVISORY COMMITTEE:

Chairperson: Prof. Eiji Oki

Member: Prof. Naoto Kishi

Member: Prof. Yasushi Yamao

Member: Prof. Minoru Terada

Member: Prof. Nattapon Kitsuwan

Date of the defense: 2017/2/01
Abstract

As Internet traffic grows with little revenue for service providers, keeping the same service level agreements (SLAs) with limited capital expenditures (CAPEX) is challenging. Major internet service providers backbone traffic grows exponentially while revenue grows logarithmically. Under such situation, CAPEX reduction and an improvement of the infrastructure utilization efficiency are both needed.

Link failures are common in Internet Protocol (IP) backbone networks and are an impediment to meeting required quality of service (QoS). After failure occurs, affected traffic is rerouted to adjacent links. This increase in network congestion leads to a reduction of addable traffic and sometimes an increase in packet drop rate. In this thesis network congestion refers to the highest link utilization over all the links in the network. An increase of network congestion may disrupt services with critical SLAs as allowable traffic becomes restricted and packet drop rate increases. Therefore, from a network operator point of view keeping a manageable congestion even under failure is desired.

A possible approach to deal with congestion increase is to augment link capacity until meeting the manageable congestion threshold. However CAPEX reduction is required. Therefore a minimization of the additional capacity is necessary. In IP networks where OSPF is widely used as a routing protocol, traffic paths are determined by link weights which are preconfigured in advance. Since traffic paths are decided by link weights, links weights therefore decide the links that will get congested. As result they determine the network congestion.
Link weights can be optimized in order to minimize the additional capacity under the worst case failure. The worst case failure is the link failure case which generates the highest congestion in the network.

In the basic model of link weight optimization, a preventive start-time optimization (PSO) scheme that determines a link weight set to minimize the worst congestion under any single-link failure was presented. Unfortunately, when there is no link failure, that link weight set leads to a congestion that may be higher than the manageable congestion. This is a penalty that will be carried on and thus become a burden especially in networks with few failures.

The first part of this thesis proposes a penalty-aware (PA) model that determines a link weight set which reduces that penalty while also reducing the worst congestion by considering both failure and non-failure scenarios. In our PA model we present two simple and effective schemes: preventive start-time optimization without penalty (PSO-NP) and strengthen preventive start-time optimization (S-PSO). PSO-NP suppresses the penalty for the no failure case while reducing the worst congestion under failure, S-PSO minimizes the worst congestion under failure and tries to minimize the penalty compared to PSO for the no failure case. Simulation results show that in several networks, PSO-NP and S-PSO achieve substantial penalty reduction while showing a congestion closed to that of PSO under worst case failure.

Despite these facts, PSO-NP and S-PSO do not guarantee an improvement of both the penalty and the worst congestion at the same time as they focus on fixed optimization conditions which restrict the emergence of upgraded solutions for that purpose. A relaxation of these fixed conditions may give us sub-optimal link weight sets that reduce the worst congestion under failure to nearly match that of PSO with a controlled penalty for the no failure case.

To determine these sub-optimal sets we expand the penalty-aware model of link weight optimization. We design a scheme where the network operator can set a manageable penalty and find the link
weight set that reduces most the worst congestion while maintaining
the penalty. This enable network operators to choose more flexible
link weight sets accordingly to their requirements under failure and
non-failure scenarios. Since setting the penalty to zero would give the
same results as PSO-NP, and not setting any penalty condition would
give S-PSO, this scheme covers PSO-NP and S-PSO. For this rea-
son we denote it: general preventive start-time optimization (GPSO).
Simulation results show that GPSO determines link weight sets with
worst congestion reduction equivalent to that of PSO under reduced
penalty for the no failure case. GPSO is effective in finding a link
weight set that reduces the congestion under both failure and non
failure cases. However it does not guarantee the manageable conges-
tion as it considers penalty.

In the second part of this thesis we propose a link-duplication
(LD) model that aims to suppress link failure in the first place in
order to always meet the manageable congestion. For this purpose
we consider the duplication or reinforcement of links which is broadly
used to make network reliable.

Link duplication provides fast recovery as only switching from
the failed link to the backup link will hide the failure at upper layers.
However, due to capital expenditure constraints, every link cannot be
duplicated. Giving priority to some selected links makes sense. As
mentioned above, traffic routes are determined by link weights that
are configured in advance. Therefore, choosing an appropriate set of
link weights may reduce the number of links that actually need to
be duplicated in order to keep a manageable congestion under any
single-link failure scenario. Now, PSO also determines the link failure
which creates the worst congestion after failure. Since by duplicating
this link we can assume it no more fails, PSO can be used to find the
smallest number of links to protect so as to guarantee a manageable
congestion under any single link failure. The LD model considers
multiple protection scenarios before optimizing link weights for the
reduction of the overall number of protected links with the congestion of keeping the congestion below the manageable threshold. Simulation results show the LD model delivers a link weight set that requires few link protections to keep the manageable congestion under any single-link failure scenario at the cost of a computation time order $L$ times that of PSO. $L$ represents the number of links in the network.

Since the LD model considers additional resources, a fair comparison with the PA model would require considering additional capacity in the PA mode as well. In the third part of this thesis we incorporate additional capacity in the PA model. For the PA model we introduce a mathematical formulation that aims to determine the minimal additional capacity to provide in order to maintain the manageable congestion under any single-link failure scenario. We then compare the LD model to the PA model that incorporates additional capacity features. Evaluation results show that the performance difference between the LD model and the PA model in terms of the required additional capacity depends on the network characteristics. The requirements of latency and continuity for traffic and geographical restriction of services should be taken into consideration when deciding which model to use.
Acknowledgements

This thesis is the summary of my doctoral study at the University of Electro-Communications, Tokyo, Japan. I am grateful to a large number of people who have helped me to accomplish this work.

First of all, I would like to express my sincere gratitude to my advisor, Professor Eiji Oki, for his mentorship, guidance and encouragements. By working under the supervision of Professor Oki, I received an invaluable experience that helped me to shape my academic and professional skills. Second, I would like to thank Tahara Hiroki for his strong assistance in collecting the data that help shape this thesis and Ihsen Aziz Ouédraogo for all his pertinent comments that greatly improved this work.

I would also like to thank Professor Professor Naoto Kishi, Professor Yasushi Yamao, Professor Minoru Terada and Professor Natapong Kitsuwan for being part of my judging committee and providing their precious comments to improve my thesis.

Finally, I want to deeply thank my parents including Mrs. Teruko Aoki for their continue financial and moral support.
Contents

List of Figures v

List of Tables vii

1 Introduction 1

1.1 Routing protocol in Internet-protocol networks . . . . . . . . . . 1
1.2 Link weight optimization . . . . . . . . . . . . . . . . . . . . . . . 3
1.3 Problem statement . . . . . . . . . . . . . . . . . . . . . . . . . . 5
1.4 Contributions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6
1.5 Organization of thesis . . . . . . . . . . . . . . . . . . . . . . . . 8

2 Basic model of link weight optimization 11

2.1 Network characteristics . . . . . . . . . . . . . . . . . . . . . . . . 11
2.2 SO: start-time optimization . . . . . . . . . . . . . . . . . . . . . 12

3 Penalty-aware model for link weight optimization 21

3.1 Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21
3.2 PSO-NP: preventive start-time with no penalty . . . . . . . . . . 22
    3.2.1 Heuristic procedure . . . . . . . . . . . . . . . . . . . . . 24
    3.2.1.1 Determination of $W^k_{SO} \in S$ . . . . . . . . . . . . . 25
    3.2.2 Computation time complexity . . . . . . . . . . . . . . . 27
    3.2.3 Performance evaluation . . . . . . . . . . . . . . . . . . . 28
3.3 S-PSO: strengthened preventive start-time optimization . . . . . . 32
    3.3.1 Heuristic procedure . . . . . . . . . . . . . . . . . . . . . 34
    3.3.2 Performance evaluation . . . . . . . . . . . . . . . . . . . 34
# List of Figures

1.1 Router and routing table .......................... 2
1.2 Composition of the thesis ........................ 9

2.1 Illustrative network to explain SO ............... 15
2.2 Traffic flows and congestion ratio with SO under no failure .... 16
2.3 Traffic flows and congestion ratio with SO under failure ....... 16
2.4 Traffic flows and congestion ratio with PSO under failure ...... 18

3.1 Technical map of discussed preventive start-time optimization schemes. 22
3.2 Generation of $W_{SO}^k \in S.$ ......................... 25
3.3 Examined network topologies ....................... 29

4.1 Relationship between worst congestion reduction ratio and upper-
bound penalty under no failure ........................ 41

5.1 Link protection Illustrative example. ............... 50
LIST OF FIGURES
List of Tables

3.1 characteristics of networks ........................................ 29
3.2 Comparisons of PSO-NP and PSO-P performance from heuristic approach .................................................... 30
3.3 Dependency of PSO-NP over $I_{\text{max}}$ ........................................ 31
3.4 Computation time comparison of SO, PSO-NP and PSO-P in seconds .................................................... 32
3.5 Comparisons of three schemes under failure and no failure scenarios 35

5.1 Protection ratios, $P(\gamma = \theta \gamma_{\text{max}})$ and $\gamma_{\text{max}}$. ........................................ 48
5.2 $\gamma$ values and protections ratios in network 1. ............... 53
5.3 $\gamma$ values and protections ratios in network 2. ............... 54
5.4 $\gamma$ values and protections ratios in network 3. ............... 54
5.5 $\gamma$ values and protections ratios in network 4. ............... 55
5.6 Comparison of computational complexities. ..................... 57

6.1 Additional resource ratios for $\gamma$ in network 1. ............... 61
6.2 Additional resource ratios for $\gamma$ in network 2. ............... 62
6.3 Additional resource ratios for $\gamma$ in network 3. ............... 62
6.4 Additional resource ratios for $\gamma$ in network 4. ............... 63
Chapter 1

Introduction

1.1 Routing protocol in Internet-protocol networks

In Internet Protocol (IP) network, a routing protocol is a set of rules that each router obeys during the process of selecting paths when sending packets from a source to a destination.

In non IP networks the same equivalent protocol exist. For example in a transportation network, cars and pedestrians must stop when lights turn red, go when green and should be prepared to stop when yellow. All theses rules constitute the routing protocol within the transportation network. The same correspondent exist and are not limited to electrical networks, water networks when sending electricity and water between a source destination pair respectively. An analogy between IP routing and Postal service is presented in [3]

In IP networks, routing protocol information or routing information are shared between routers throughout the network. These routing information are composed of entries that are stored in a so called routing table(See Fig 1.1). Each entry is a couple composed of a destination address and an output port. When a packet arrives at a router, the packet destination address is searched inside the routing table. If there is a match, the packet will be forward through the output port that corresponds to the matching destination address.
1. INTRODUCTION

These entries (destination address and output ports) can either be set manually or dynamically. In small networks the network operator may manually set the entries of each router in the network [42]. However, in larger networks which comprise tens or hundreds of routers, setting each router is troublesome. Also, link failure becomes more recurrent and will require to reset all the affected routers making it harder to use manual configuration of entries in large IP networks.

In large networks where dynamical configuration is employed, OSPF (Open Shortest Path First) [48] is widely used as a protocol to set and reset entries at routers by taking into consideration network topology and failure conditions. In OSPF each router shares its routing information with the rest of routers of the network within a given period in order to build a network connectivity entries. When failure happens on a particular link, the adjacent routers detect it as OSPF keep alive packets will no more be received. These adjacent routers will then change their entries in order to be able to communicate again by using detour routes obtained through information previously shared via OSPF.

Now, for a given destination address the number of possibilities to send pack-
ets can be enormous when the number of router grows. This means the same destination can be associated with different output ports at the same router. As storage as well as search speed will become a challenge differentiation is required.

To select specific routes in OSPF, each link is allocated a link weight or link cost which is set at each router interface. Based on these link weights the Dijkstra algorithm [9] computes the shortest path between any given node to the rest of the nodes in the network. For a given source and destination pair, the shortest path is the one that has the lowest cost among all the paths between the same pair. Each node keeps the shortest path between itself and others and relies on that list of entries to send packet to the adjacent node situated on the suitable shortest path.

Since the capacity of a link to carry data is finite an increase in traffic augments link utilization or congestion leading to packet lost and packet transmission delay. A possible way to deal with this situation is to increase link capacity to the point where congestion becomes lower than the manageable threshold. With the continued increase of traffic demand and the decrease of data unit price network operators can afford to always increase capacity. Therefore a reduction of the additional capacity under high traffic demand is required.

Since traffic path are determined by link weights which can be changed by network operators, link weights can be tuned to divert or deflect traffic in order to relax congested paths [53] or links thus creating room for more traffic under a manageable congestion level.

1.2 Link weight optimization

Start-time optimization (SO) is used to determine the suitable link weight set at the beginning of network operation in order to minimize congestion. Multiple algorithms that focus on implementing SO are presented in [15, 22, 23, 65, 68]. These schemes all assume that, the network topology and the traffic volume from node-2-node is given. A tabu search based heuristic algorithm is discussed in [22, 23]. Buriol et al. presented a genetic algorithm with a local improvement procedure [15]. A fast heuristic algorithm was also developed in [65]. These link
weight optimization algorithms provide nearly an optimal solution in a practical manner.

Unfortunately, SO cannot deal with link failures. In OSPF, when a link failure occurs, traffic is rerouted to other output ports making the corresponding links more congested. As congestion in these links trespasses manageable levels, SO is no more suitable under a failure scenario. An approach to tackle this drawback would be to re-optimize link weights after a link failure occurs. However, this scheme leads to network instability [4] as 50% of link failures last less than one minute in some networks [27]. Moreover, for different causes such as maintenance activities, link failures are common [41]. Therefore, a scheme that considers any possible failure in advance to determine link weights in order to keep the manageable congestion even under failure is required.

Considering link failures in advance to determine a suitable link weight set against failure is studied in the preventive start-time optimization (PSO). There are many schemes that were presented in the family of PSO. In [29], only single link failures are considered as 70% of link failures are single ones [27, 41]. Other works that focus on PSO [30, 63, 64] have been introduced. The first PSO related schemes were respectively, PSO-LC (limited range of candidates) [29] and PSO-WC (wide range of candidates) [30]. Even though PSO-LC reduces the congestion ratio under the worst-case link failure, it does not confirm the optimal worst-case performance. PSO-WC improves PSO-LC by upgrading the objective function of SO. It considers all possible single link failures at start time to determine the set of link weights that minimizes the worst-case congestion. Numerical analysis showed that PSO-WC effectively relaxes the worst-case network congestion compared to SO, while avoiding network instability as run-time changes of link weights is not performed.

Also, PSO-WC performance is superior to that of PSO-LC [30]. Ranaweera et al. presented a PSO policy for the hose model, where the exact traffic demand between each source and destination node pair does not need to be specified, to optimize the link weights against link failures [63, 64]. Their presented scheme for the hose model employs a heuristic algorithm to determine a suitable set of link weights to reduce the worst-case congestion for any single link failure.
1.3 Problem statement

The authors in [29] pointed out that under no failure, the congestion shown by the PSO is higher than that of SO, but they do not show how to decrease the congestion. This is a penalty especially for networks with few failures. This penalty will be a burden in the long run since we do not change link weights during operation time as it would cause network instability. Now let us call this preventive start-time optimization scheme in [29], PSO-P (preventive start-time with penalty). A question arises: Is it possible to find a link weight set that eliminates that penalty through penalty-aware (PA) model for link weight optimization?

We could also think of network where failure do not impact the routing layer, in the first place. To avoid link failure effect at the routing layer link duplication or link redundancy [19, 36] is used for reliability. It provides redundancy of links and only switching automatically from the failure affected link to the backup link is enough to provide instantaneous recovery and hide the failure impact at the routing layer. From a network operator perspective it presents advantages in terms of operation and fast convergence under failure. This is the main reason why it is broadly used in large commercial networks.

Since the possibility of having multiple link failures at the same time is low [27], the network operator can assume that a duplicated link will not fail from the IP layer’s point of view.

However, in a link-duplication (LD) model where links are reinforced doubling each links in the network requires a lot of resources which will increase capital expenditures. Therefore, keeping the same QoS while reducing the number of protected links is necessary. Since traffic paths are decided by link weights, links weights therefore determine the links that need to be duplicated. The second question arises: How can we optimize link weights so as to minimize the number of links to duplicate?

If we had considered additional resources in the PA model, the third question arises: how much resources would be needed to always keep the manageable congestion? The last question is: Between the PA model and the LD model which approach is better from a network operator point of view?
1. INTRODUCTION

1.4 Contributions

We propose a penalty-aware (PA) model for link weights optimization to tackle the penalty of PSO under no failure. The PA model includes two schemes: preventive start-time with no penalty (PSO-NP) [34] and strengthened preventive start-time optimization (S-PSO) [32]. PSO-NP generates a link weight set that completely suppresses the penalty and at the same time, reduces substantially the congestion even for the worst case congestion scenario. PSO-NP is based on SO. SO determines multiple sets that minimize the congestion under a non failure scenario, but the firstly found set is chosen as solution. PSO-NP evaluates the performance of each of these sets under the worst failure and chooses the one that reduces most the congestion. On the contrary S-PSO consider multiple sets that are solutions found by PSO and chooses the one with lowest penalty. In that sense PSO-NP minimizes the penalty and reduce the worst congestion while S-PSO minimizes the worst congestion but reduce the penalty. We present both heuristic procedure and mathematical formulations for PSO-NP and S-SPO. For each of these schemes performance evaluation and comparison with PSO-P is made. A complete analysis of the computation time complexity is provided as well. Simulation results show that PSO-NP and S-PSO achieve substantial congestion reduction for any failure case while dealing with the penalty in case of no failure in multiple network topologies.

However, PSO-NP and S-PSO cannot always provide improvement as they find a link weight set that is element of the sets determined by SO and PSO respectively. To tackle this deficiency we consider to determine sub-optimal link weight sets with marginal deviation from both the penalty and minimal worst congestion. In other words we consider the trade-off between the penalty when there is no failure and the reduction of the worst case congestion. We explore this in a generalized penalty-aware (GPA) model for link weight optimization. In that model we presented a Generalized preventive start-time optimization (GPSO) [51]. GPSO determines the weight set that minimizes the worst case congestion ratio when the allowable penalty is given. As the allowable penalty can be tuned, GPSO covers PSO-NP and S-PSO. It is therefore an expansion of both PSO-NP and S-PSO. We also conduct an evaluation of GPSO performance after presenting
its mathematical formulation and heuristic procedure. Evaluation results show that GPSO finds link weight sets with marginal penalty and worst congestion equivalent to that of PSO-P.

Concerning the second question, we propose a link-duplication (LD) model for link weight optimization in IP networks. This model is based on PSO. PSO considers any failure patterns to determine at the start time a link weight set that minimizes the worst congestion. PSO also determines the link failure which creates the worst congestion after failure. Therefore PSO can be used to find the suitable link to duplicate in order to tackle link failure.

In the LD model, we present a scheme that uses PSO to reduce the number of links to duplicate at the start time of network operation with the condition of keeping a given manageable congestion under any single-link failure [31]. However, the links selected in [31] are chosen after PSO determined link weights are set. As the selected links reflects the obtained link weights, some protection patterns that can guarantee the pre-defined manageable congestion with a lower number of protected links might be skipped. This could lead to over-provision of resources. We develop a new scheme in [33]. We consider each link reinforcement scenario before optimizing link weights. During the optimization procedure, links that are reinforced are no more considered as failure targets as we assume that reinforced links do not fail. The best reinforcement is the reinforcement scenario where the optimized link weight shows the lowest congestion under the worst case link failure. Simulation results show that, to keep the same manageable congestion under failure, the newly develop scheme reduces the number of links to protect compared to the scheme in [31].

Finally, since the LD model considers additional resources, a fair comparison with the PA model would require considering additional capacity in the PA model as well. To answer the third question, we incorporate additional capacity in the PA model. For the PA model we introduce a mathematical formulation that aims to determine the minimal additional capacity to provide in order to maintain the manageable congestion under any single-link failure scenario. We answer the fourth question by comparing the LD model to the PA model that incorporates additional capacity features. Evaluation results show that the performance difference between the LD model and the PA model in terms of the required
1. INTRODUCTION

additional capacity depends on the network characteristics. The requirements of latency and continuity for traffic and geographical restriction of services should be taken into consideration when deciding which model to use.

1.5 Organization of thesis

The organization of the thesis is shown in Fig. 1.2. The Basic model of link weight optimization that includes conventional link weight optimization schemes such as SO and PSO is presented in Chapter 2. The PA model which includes PSO-NP and S-PSO is discussed in Chapter 3. GPA model which contains GPSO and covers PSO-NP and S-PSO and aims to balance the congestion under failure and non failure cases, follows in Chapter 4. Chapter 5 includes the LD model and its related link weight optimization techniques for link duplication. Chapter 6 compares the LD model and the PA model that considers additional capacity. Chapter 7 concludes the thesis by summarizing the findings and showing the direction for future work.
1.5 Organization of thesis

Figure 1.2: Composition of the thesis
1. INTRODUCTION
Chapter 2

Basic model of link weight optimization

In this chapter we explain a basic link weight optimization model under single link failure scenario where additional capacity is not considered. In other words, a link weight optimization model to minimize the network congestion ratio. In Section 2.1, network characteristics considered in this thesis is mentioned. Details of SO which does not consider link failure when minimizing the congestion is presented in Section 2.2. Section 2.3 presents PSO which determines the suitable link weights in order to minimize network congestion that arises from any single-link failure.

2.1 Network characteristics

The directed graph $G(V, E)$ represents the network, where $V$ is the set of nodes and $E$ is the set of links. The number of nodes is $N = |V|$. An individual node is represented as $v \in V$, where $v = 1, 2, \cdots, N$. We consider only single link failures in this work, as the probability of having simultaneous link failures is much less than that of a single link failure [27, 38, 41, 62, 70]. We all also consider that each link has both directions. Let a link from node $i \in V$ to node $j \in V$ be denoted as $(i, j) \in E$. $L$ is the number of links in the network, or $L = |E|$. $F = \{0, 1, 2, \ldots, L\}$ is the set of link failure indices. For $l \in F$, $l = 0$ indicates no link failure and $l(\neq 0)$ indicates the failure of $(i, j) = l(\neq 0)$. The network in which
2. BASIC MODEL OF LINK WEIGHT OPTIMIZATION

link \( l(\neq 0) \) is considered failed is described as a directed graph \( G_l(V, E_l) \). Since \( l = 0 \) indicates no link failure, \( G_0(V, E_0) = G(V, E) \). \( u_{ij}, (i, j) \in E \) represents the traffic flowing through link \((i, j)\) and \( c_{ij} \) its capacity. If \((i, j) = l\) and \( l \) failed then \( c'_{ij} = 0 \). \( W = \{w_{ij}\} \) is the link weight set of network \( G \), where \( w_{ij} \) is the weight of \((i, j)\). Let \( \{1, \ldots, w_{\text{max}}\} \) be the set of values possibly taken as link weights. \( x_{ij}^p(W,l) \) is the portion of traffic from node \( p \in V \) to node \( q \in V \) routed through \((i, j) \in E_l \). Note that \( x_{ij}^p(W,l) \) will be determined based on the shortest path routing when link weight set \( W \) is applied to the network. In this analysis, it is assumed that Equal-Cost Multi-Path (ECMP) routing is employed, where traffic is evenly split among equal-cost paths \[67\]. \( x_{ij}^p(W,l) \) is used to represent the load distribution variables under link weights set \( W \). A traffic matrix \( T \) is defined by \( T = \{d_{pq}\} \), where \( d_{pq} \) represents the traffic rate from node \( p \) to \( q \).

Let consider \( E(W) \) the set of links on our transmitting paths when \( W \) is our link weight set. The network congestion ratio that we denote as \( r \), represents the maximum value of all link utilization ratios in the network.

\[
r(W) = \max_{(i,j) \in E(W)} \frac{u_{ij}}{c_{ij}},
\]

where \( 0 \leq r(W) \leq 1 \). Additional traffic can be maximized by choosing the suitable weight set \( W \) which minimizes \( r(W) \) \[53\]. For any link \( l \in F \) failed, the congestion ratio is defined as:

\[
r(W,l) = \max_{(i,j) \in E_l(W)} \frac{u_{ij}}{c_{ij}},
\]

and the worst case congestion ratio as:

\[
R(W) = \max_{l \in F} r(W,l).
\]

Note that \( r(W,0) \) represents the congestion ratio in a network when there is no failure. For simplicity, let \( r(W,l) \) be denoted as \( r(l) \). \( r(W,0) \) is therefore \( r(0) \).

2.2 SO: start-time optimization

SO determines an optimal link weight set \( W \) that minimizes \( r(0) \) when link failure is not considered, enabling additional flow in the network without extra capacity.
2.2 SO: start-time optimization

Let us call that set $W_{SO}$. It is expressed as:

$$W_{SO} = \arg \min_{W} r(0). \quad (2.4)$$

$r(W_{SO}, l)$ is denoted as $r_{SO}(l)$. When the network topology and the traffic demands are known, $W$ that minimizes $r(0)$ is obtained through a integer linear programming (ILP) formulation presented in [18]. The transformation of an optimization problem into an ILP expression is useful as it reduces the complexity of the steps to reach the optimal solution. This gives the possibility to find an optimal solution for cases where the size of the parameters is small. These solutions will then help to check the validity of heuristic methods that deliver nearly optimal solution as the ILP becomes hard to solve when the size of the size of the parameters becomes large. The following ILP represents that of SO and is expressed as follows:

Objective  \[ \min_{W} r(W, 0) \quad (2.5a) \]

s.t.  \[ \sum_{j:(i,j) \in E_l} x_{ij}^{pq}(W, l) - \sum_{j:(i,j) \in E_l} x_{ji}^{pq}(W, l) = 1, \quad \forall p, q \in V, i = p, l \in F \quad (2.5b) \]

\[ \sum_{j:(i,j) \in E_l} x_{ij}^{pq}(W, l) - \sum_{j:(i,j) \in E_l} x_{ji}^{pq}(W, l) = 0, \quad \forall p, q \in V, i(\neq p, q), l \in F \quad (2.5c) \]

\[ \sum_{j:(i,j) \in E_l} x_{ij}^{pq}(W, l) - \sum_{j:(i,j) \in E_l} x_{ji}^{pq}(W, l) = -1, \quad \forall p, q \in V, i = q, l \in F \quad (2.5d) \]

\[ \sum_{p,q \in V} d_{pq} x_{ij}^{pq}(W, 0) \leq c_{ij}^0 \cdot r_{SO}, \quad \forall (i, j) \in E_0 \quad (2.5e) \]

\[ \sum_{p,q \in V} d_{pq} x_{ij}^{pq}(W, l) \leq c_{ij}^l \cdot r_{PSO}, \quad \forall (i, j) \in E_l, l \in F \quad (2.5f) \]

\[ 0 \leq f_{pq}^l(l) - x_{ij}^{pq}(W, l) \leq 1 - \delta_{ij}^q(l), \forall p, q \in V, (i, j) \in E_l, l \in F \quad (2.5g) \]
2. BASIC MODEL OF LINK WEIGHT OPTIMIZATION

\[ x_{ij}^p(W,l) \leq \delta_q^j(l), \forall p, q \in V, \]
\[ (i,j) \in E_l, l \in F \] (2.5h)
\[ 0 \leq \psi_{jq}(l) + w_{ij} - \psi_{iq}(l) \leq (1 - \delta_q^j(l))U, \]
\[ \forall q \in V, (i,j) \in E_l, l \in F \] (2.5i)
\[ 1 - \delta_q^j(l) \leq \psi_{jq}(l) + w_{ij} \leq \psi_{iq}(l), \]
\[ \forall q \in V, (i,j) \in E_l, l \in F \] (2.5j)
\[ f_{pq}^i(l) \geq 0, \forall p, q, i \in V, l \in F \] (2.5k)
\[ \delta_q^j(l) \in \{0, 1\}, \forall q \in V, (i,j) \in E_l, l \in F \] (2.5l)
\[ 1 \leq w_{ij} \leq w_{max}, \forall (i,j) \in E_l, l \in F. \] (2.5m)

The key decision variable of this ILP problem is the link weight set \( W \). Other decision variables are \( f_{pq}^i(l), x_{ij}^p(W,l), \delta_q^j(l), \delta_q^i \) and \( \psi_{jq}(l) \). These variables are determined once \( W \) is calculated.

The constraints of Eqs. (2.5b)-(2.5m) are explained as follows. Eqs. (2.5b)-(3.3c) express the flow conservation constraints. Eq. (2.5e) expresses the capacity constraint for each link by using the network congestion ratio, \( r_{SO} \), in case of no link failure. Eq. (2.5f) expresses the capacity constraint for each link by using the worst-case network congestion ratio, \( r_{PSO} \), for any single link failure. Eqs. (2.5g) - (2.5j) indicate the constraints of the ECMP routing. Eq. (2.5g) indicates constraints of equal traffic splitting at node \( i \). \( f_{pq}^i(l) \) represents the traffic portion from node \( p \) to \( q \) that arrives at node \( i \) under \( l \in F \), which is equally divided by the number of outgoing links from node \( i \) that are on the shortest path to \( q \). \( \delta_q^j(l) \) is a binary variable which is 1 if \( (i,j) \) is on the shortest path to \( q \) under \( l \in F \) and otherwise 0. Eq. (2.5g) forces all the non-zero outgoing traffic portions from node \( i \), \( x_{ij}^p(W,l) \) for \( (i,j) \in E_l \), on the shortest paths to be equal to \( f_{pq}^i(l) \); all the non-zero \( x_{ij}^p(W,l) \) from node \( i \) are equal. Eq. (4.1h) indicates the upper bound of \( x_{ij}^p(W,l) \) in the ECMP. If \( (i,j) \) is on the shortest path, then \( x_{ij}^p(W,l) \) is at most 1. As \( \delta_q^j(l) = 1 \), Eq. (4.1h) is verified. Otherwise, \( (i,j) \) is not on the shortest path and \( x_{ij}^p(W,l) = \delta_q^j(l) = 0 \), which also verifies Eq. (4.1h). Eqs. (2.5i) and (2.5j) express link cost constraints of \( (i,j) \) in the ECMP, where \( \psi_{jq}(l) \) is the cost of the shortest path from \( j \) to \( q \) under \( l \in F \). If \( (i,j) \) is on the shortest path from \( p \) to \( q \), \( \psi_{jq}(l) + w_{ij} - \psi_{iq}(l) = 0 \), which verifies Eq. (2.5i).
2.2 SO: start-time optimization

Otherwise, $\psi^{jq}(l) + w_{ij} - \psi^{iq}(l) \geq 1$ as $w_{ij} \geq 1$; since $\delta_{jq}^{ij}(l) = 0$, Eq. (2.5j) is verified. Note that $U$ is a given constant with a sufficiently large value to verify Eq. (2.5i) when $(i, j)$ is not on the shortest path. The ECMP routing constraints are explained in more details in [18, 60]. Eqs. (3.3k)-(2.5m) give the types and ranges of the decision variables.

Let us consider the illustrative network at Fig. 2.1. An operator wants to send 1 unit of traffic from respectively nodes $v_1$, $v_2$ to node $v_4$ with all link capacities uniformly set to 1.

![Illustrative network to explain SO](image)

**Figure 2.1:** Illustrative network to explain SO

These traffic requirements and capacity conditions are assigned to the above ILP. The resulting weight set is shown in Fig. 2.2. Under this weight set the traffic requirements from respectively $v_1$ to $v_4$ and $v_2$ to $v_4$ are sent such that the congestion ratio becomes equal to 0.75.

Unfortunately, SO is not optimal under failure. Under failure, traffic is rerouted to other non failed links. This increases the congestion ratio of the network. Under the worst case failure especially, we have:

$$R(W_{SO}) \geq r_{SO}(0). \quad (2.6)$$

For higher values of $R(W_{SO})$ data loss may occur and lead to a decrease in quality of service.

When the diagonal link fails with the traffic flows described in Fig. 2.2, traffic is rerouted as shown in Fig. 2.3 and the congestion ratio rises to 1.5 confirming Eq. (2.6). To tackle this drawback, PSO-P [29] was introduced.
2. BASIC MODEL OF LINK WEIGHT OPTIMIZATION

![Traffic flows and congestion ratio with SO under no failure.](image1)

**Figure 2.2:** Traffic flows and congestion ratio with SO under no failure.

![Traffic flows and congestion ratio with SO under failure.](image2)

**Figure 2.3:** Traffic flows and congestion ratio with SO under failure.

2.3 PSO-P: preventive start-time optimization with penalty

PSO-P considers any possible failure pattern to determine at the start time a link weight set that minimizes the worst congestion ratio. The main achievement is the preservation of network stability, because there are no running-time changes. Since PSO-P focuses only on finding a link weight set that minimizes $R(W)$, a
solution is expressed as:

$$W_{PSO-P} = \arg \min_W R(W).$$  \hfill (2.7)

An ILP formulation of PSO-P is deduced from that of SO and expressed in the following mathematical formulation.

Objective

$$\min_W R(W)$$ \hfill (2.8a)

subject to

$$\sum_{j:(i,j) \in E_l} x_{ij}^{pq} (W,l) - \sum_{j:(i,j) \in E_l} x_{ji}^{pq} (W,l) = 1,$$

$$\forall p, q \in V, i = p, l \in F$$ \hfill (2.8b)

$$\sum_{j:(i,j) \in E_l} x_{ij}^{pq} (W,l) - \sum_{j:(i,j) \in E_l} x_{ji}^{pq} (W,l) = 0,$$

$$\forall p, q \in V, i \neq p, q, l \in F$$ \hfill (2.8c)

$$\sum_{j:(i,j) \in E_l} x_{ij}^{pq} (W,l) - \sum_{j:(i,j) \in E_l} x_{ji}^{pq} (W,l) = -1,$$

$$\forall p, q \in V, i = q, l \in F$$ \hfill (2.8d)

$$\sum_{p,q \in V} d_{pq} x_{ij}^{pq} (W,0) \leq c_{ij}^0 \cdot r_{SO},$$

$$\forall (i,j) \in E_0$$ \hfill (2.8e)

$$\sum_{p,q \in V} d_{pq} x_{ij}^{pq} (W,l) \leq c_{ij}^l \cdot r_{PSO},$$

$$\forall (i,j) \in E_l, l \in F$$ \hfill (2.8f)

$$0 \leq f_{pq} (l) - x_{ij}^{pq} (W,l) \leq 1 - \delta_{i,j}^q (l), \forall p, q \in V,$$

$$(i,j) \in E_l, l \in F$$ \hfill (2.8g)

$$x_{ij}^{pq} (W,l) \leq \delta_{i,j}^q (l), \forall p, q \in V,$$

$$(i,j) \in E_l, l \in F$$ \hfill (2.8h)

$$0 \leq \psi_i^q (l) + w_{ij} - \psi_i^q (l) \leq (1 - \delta_{i,j}^q (l)) U,$$

$$\forall q \in V, (i,j) \in E_l, l \in F$$ \hfill (2.8i)

$$1 - \delta_{i,j}^q (l) \leq \psi_i^q (l) + w_{ij} \leq \psi_i^q (l),$$

$$\forall q \in V, (i,j) \in E_l, l \in F$$ \hfill (2.8j)

$$f_{pq}^i (l) \geq 0, \forall p, q, i \in V, l \in F$$ \hfill (2.8k)
2. BASIC MODEL OF LINK WEIGHT OPTIMIZATION

\[ \delta_{ij}(l) \in \{0, 1\}, \forall q \in V, (i, j) \in E_l, l \in F \quad (2.8i) \]
\[ 1 \leq w_{ij} \leq w_{\text{max}}, \forall (i, j) \in E_l, l \in F. \quad (2.8m) \]

Generally the worst case congestion in SO is higher or equal to that of PSO-P, which is expressed by:

\[ R(W_{SO}) \geq R(W_{PSO-P}). \quad (2.9) \]

In other words, PSO-P outperforms SO under the worst case congestion with a significant worst congestion ratio reduction. We define this reduction ratio of PSO-P as:

\[ \alpha_{PSO-P} = \frac{R(W_{SO}) - R(W_{PSO-P})}{R(W_{SO})}. \quad (2.10) \]

If we had considered PSO-P to determine link weights for the illustrative network introduced at Fig. 2.1, traffic flows would be as shown in Fig. 2.4 and the network worst case congestion ratio \( R(W_{PSO-P}) \) be equal to 1.0. By replacing the values of \( R(W_{PSO-P}) \) and \( R(W_{SO}) = 1 \) in Eq. (2.10) we have reduction ratio \( \alpha_{PSO-P} = 0.5 \) meaning 50\% reduction compared to the congestion ratio with SO under failure.

\[ \frac{\text{max}_{i \in \delta} r_{\max}(i) - \text{max}_{i \in \delta} r_{\text{PSO}}(i)}{\text{max}_{i \in \delta} r_{\max}(i)} = \frac{15 - 1.0}{1.0} = 0.5 \]

**Figure 2.4:** Traffic flows and congestion ratio with PSO under failure.

The difficulty with PSO-P is that if the link weight set determined by PSO-P is used in a no failure scenario, the congestion ratio may be higher than that of
2.3 PSO-P: preventive start-time optimization with penalty

In general, under no failure the congestion ratio of SO is less or equal to that of PSO-P:

\[ r(W_{SO}, 0) \leq r(W_{PSO-P}, 0). \] (2.11)

Therefore, PSO-P shows a penalty under no failure that we define as:

\[ \beta_{PSO-P} = \frac{r(W_{PSO-P}, 0) - r(W_{SO}, 0)}{r(W_{SO}, 0)}. \] (2.12)

In the illustrative example, \( r(W_{PSO-P}, 0) = 1.0 \). Since \( r(W_{SO}, 0) \) is equal to 0.75 the penalty \( \beta_{PSO-P} = 0.33 \). This means that PSO-P shows a congestion ratio 33% higher than that of SO.

\( \beta_{PSO-P} \) raises an issue for networks with relatively few link failures because the penalty is carried on and becomes a burden in the long run. We address this issue in the next chapter by proposing a penalty-aware model of link weight optimization. This model is composed of two schemes. The first one is preventive start-time with no penalty (PSO-NP) [34] which eliminates the penalty by keeping \( \beta = 0 \) and increases \( \alpha \). The second one is strengthened preventive start-time optimization (S-PSO) [32] which keeps the same \( \alpha \) as PSO-P and reduces \( \beta \) under the worst case failure scenario.
2. BASIC MODEL OF LINK WEIGHT OPTIMIZATION
Chapter 3

Penalty-aware model for link weight optimization

3.1 Overview

This chapter proposes a penalty-aware (PA) model to reduce both the penalty and the worst congestion ratio without additional capacity in the network by considering both failure and non-failure cases. In this model we introduce two schemes that are respectively PSO-NP and S-PSO. We also make and evaluation of both schemes and compare them to PSO-P. Comparison parameters are $\alpha$ and $\beta$.

In Section 3.2 PSO-NP which determines a link weight set that eliminates the penalty of PSO-P compared to SO when there is no link failure is presented. PSO-NP determines a link weight set that completely suppresses the penalty and at the same time, reduces substantially the congestion ratio even for the worst case congestion scenario. PSO-NP scheme is based on SO. Under no failure scenario, SO generates a set that minimizes the congestion ratio. Many candidate sets with the same performance may exist. SO chooses the first one among these sets. PSO-NP extends SO by evaluating the performance of each of these sets under the worst-case failure to choose the one that reduces most the congestion ratio.

Unfortunately, compared to SO, PSO-NP does not guarantee congestion reduction under failure, as its generated weight set may sometimes be equal to the
3. PENALTY-AWARE MODEL FOR LINK WEIGHT OPTIMIZATION

set determined by SO. With PSO-NP, the weakness of SO under failure is not completely solved.

In section 3.3 we introduce S-PSO to solve that weakness while reducing the penalty of PSO-P. S-PSO is based on PSO. Analogically to SO, PSO-P determines the targeted weight sets which minimizes the worst congestion, but multiple targeted sets with the same minimal worst congestion ratio may exist. In [30], the targeted weight set firstly found by the conventional PSO is selected as a solution. Our presented S-PSO scheme collects all these targeted weight sets and evaluates their congestion under no failure. Among the targeted sets, the one which shows the minimum congestion under no failure is chosen as a solution. A technical map of SO, PSO-P, PSO-NP, S-PSO is shown in Fig. 3.1 where each scheme is associated with its corresponding target \((\alpha, \beta)\) couple.

3.2 PSO-NP: preventive start-time with no penalty

As the solution weight set of PSO-NP, \(W_{PSO-NP}\) is generated through SO, under no failure PSO-NP and SO show the same congestion ratio:

\[
r(W_{SO}, 0) = r(W_{PSO-NP}, 0).
\]  
(3.1)
Moreover, PSO-P minimizes the worst congestion ratio. As a result,

\[ R(W_{PSO-P}) \leq R(W_{PSO-NP}) \leq R(W_{SO}). \]  

(3.2)

PSO-NP guarantees zero penalty under no failure while boosting protection under failure compared to the typical SO scheme. An ILP formulation of PSO-NP can therefore be expressed as:

\[
\text{Objective } \min_{W} \{ r(W, 0) + \epsilon \cdot R(W) \} \quad (3.3a)
\]

subject to

\[
\sum_{j:(i,j) \in E_t} x_{ij}^{pq}(W, l) - \sum_{j:(i,j) \in E_t} x_{ji}^{pq}(W, l) = 1, \quad \forall p, q \in V, i = p, l \in F \quad (3.3b)
\]

\[
\sum_{j:(i,j) \in E_t} x_{ij}^{pq}(W, l) - \sum_{j:(i,j) \in E_t} x_{ji}^{pq}(W, l) = 0, \quad \forall p, q \in V, i(\neq p, q), l \in F \quad (3.3c)
\]

\[
\sum_{j:(i,j) \in E_t} x_{ij}^{pq}(W, l) - \sum_{j:(i,j) \in E_t} x_{ji}^{pq}(W, l) = -1, \quad \forall p, q \in V, i = q, l \in F \quad (3.3d)
\]

\[
\sum_{p,q \in V} d_{pq} x_{ij}^{pq}(W, 0) \leq c_{ij}^0 \cdot r_{SO}, \quad \forall (i, j) \in E_0 \quad (3.3e)
\]

\[
\sum_{p,q \in V} d_{pq} x_{ij}^{pq}(W, l) \leq c_{ij}^l \cdot r_{PSO}, \quad \forall (i, j) \in E_l, l \in F \quad (3.3f)
\]

\[
0 \leq f_{pq}^l - x_{ij}^{pq}(W, l) \leq 1 - \delta_{ij}^q(l), \forall p, q \in V, \quad (i, j) \in E_l, l \in F \quad (3.3g)
\]

\[
x_{ij}^{pq}(W, l) \leq \delta_{ij}^q(l), \forall p, q \in V, \quad (i, j) \in E_l, l \in F \quad (3.3h)
\]

\[
0 \leq \psi_{pq}^q(l) + w_{ij} - \psi_{ij}^q(l) \leq (1 - \delta_{ij}^q(l))U, \quad \forall q \in V, (i, j) \in E_l, l \in F \quad (3.3i)
\]

\[
1 - \delta_{ij}^q(l) \leq \psi_{pq}^q(l) + w_{ij} \leq \psi_{ij}^q(l), \quad \forall q \in V, (i, j) \in E_l, l \in F \quad (3.3j)
\]
3. PENALTY-AWARE MODEL FOR LINK WEIGHT OPTIMIZATION

\[ f_{pq}^l(l) \geq 0, \forall p, q, i \in V, l \in F \]  \hspace{1cm} (3.3k)

\[ \delta_l^i(l) \in \{0, 1\}, \forall q \in V, (i, j) \in E_l, l \in F \]  \hspace{1cm} (3.3l)

\[ 1 \leq w_{ij} \leq w_{max}, \forall (i, j) \in E_l, l \in F. \]  \hspace{1cm} (3.3m)

Where the key decision variables are defined as in SO and PSO-P formulations.

Unfortunately it is not possible to get the optimal solution when the network size becomes large. For this reason we present a heuristic algorithm to determine \( W_{PSO-NP} \) for any network topology.

3.2.1 Heuristic procedure

Since SO may generate multiple link weight set that minimizes \( r(0) \), let \( W_{SO}^k \) be such sets, where \( k \) is an index. Let us regroup these sets in a set \( S \). All these sets present the same minimal congestion ratio under no failure. Therefore, \( r(W_{SO}^k, 0) = r(W_{SO}, 0) \) for all \( W_{SO}^k \in S \). Thus the penalty for all the elements of \( S \) under no failure would be such that:

\[ \beta_k = \frac{r(W_{SO}^k, 0) - r(W_{SO}, 0)}{r(W_{SO}, 0)} = 0, \]  \hspace{1cm} (3.4)

for all \( W_{SO}^k \in S \). As all the elements of \( S \) show zero penalty, the one showing the lowest congestion ratio under the worst case failure is by definition \( W_{PSO-NP} \).

Let

\[ \alpha_k = \frac{R(W_{SO}) - R(W_{SO}^k)}{R(W_{SO})}, \]  \hspace{1cm} (3.5)

be the reduction ratio of the worst case congestion of \( W_{SO}^k \in S \). Under worst case failure, it is desirable to have a large value of \( \alpha \) as it represents the worst congestion reduction ratio. Since \( W_{PSO-NP} \) is the element of \( S \) with the lowest congestion ratio under the worst case failure, it is equivalent to the set with the largest \( \alpha \). \( W_{PSO-NP} \) is obtained by:

\[ W_{PSO-NP} \equiv \arg \max_{W_{SO}^k \in S} \alpha_k^k. \]  \hspace{1cm} (3.6)

Therefore, \( W_{PSO-NP} \) eliminates the penalty while reducing the worst congestion ratio compared to \( W_{SO} \).
The procedure taken by PSO-NP to obtain $W_{PSO-NP}$ is defined by Eq. (3.6) and divided into three steps as follows. In our procedure, $S$ is initialized to $S = \emptyset$. $w_{max}$ is set to 1000.

- **Step 1**: Get $W_{SO}$ and collect all sets that generates the same $r(W_{SO}, 0)$. These sets are elements of $S$. The method to determine these sets is mentioned in the next section.

- **Step 2**: For each $W^k_{SO} \in S$, evaluate $\alpha_k$, using Eq. (3.5).

- **Step 3**: Get $W_{PSO-NP}$ by using Eq. (3.6).

### 3.2.1.1 Determination of $W^k_{SO} \in S$.

$W^k_{SO} \in S$ is determined through a tabu search [22, 23, 29]. The concept of how to determine $W^k_{SO} \in S$ is depicted in Fig. 3.2. Here, multiple randomly generated initial weights are considered. The elements of each initial weight set $W_{init}$ follows a uniform distribution from 1 to $w_{max}$. For each every $W_{init}$ set we calculate the link weight that shows the local minima in terms of congestion ratio when there is no failure. The weight sets that show the lowest global minimum are selected as elements of $S$.

![Figure 3.2: Generation of $W^k_{SO} \in S$.](image-url)
3. PENALTY-AWARE MODEL FOR LINK WEIGHT OPTIMIZATION

Step 1 is explained in details. It is divided into 3 steps. $z$ represents a fixed increment value in the link weight searching procedure. $r$ and $I$ are variables used to respectively store the congestion ratio and count the number of used initial weight sets. These variables are respectively initialized to $r = \infty$ and $I = 1$. The procedure in step 1 is as follows:

- **Step 1.1:**
  
  Generate initial link weight set $W_{init}$ randomly as mentioned at first paragraph of section 3.2.1.1.
  
  $W_{temp} \leftarrow W_{init}$.
  
  $flag \leftarrow true$.

- **Step 1.2:**
  
  While ($flag$)
  
  Set the routing paths in the network by using the shortest path algorithm and find the link $(i_0, j_0)$ with the highest congestion ratio $r_0$.
  
  if($r_0 = r$)
    
    add $W_{temp}$ into $S$.
  
  end.
  
  if($r_0 < r$)
    
    $r \leftarrow r_0$.
    
    clear $S$.
    
    add $W_{temp}$ into $S$.
  
  end.

  Update $W_{temp}$ by replacing $w_{i_0j_0} \leftarrow w_{i_0j_0} + z$.
  
  if ($w_{i_0j_0} + z > w_{max}$)
    
    $flag \leftarrow false$.
  
  end.
3.2 PSO-NP: preventive start-time with no penalty

end.

- Step 1.3:
  Increment \( I \).
  if \( I \leq I_{\text{max}} \)
    go to step 1.1.
  else
    return \( S = \{W^1_{SO}, W^2_{SO}, W^3_{SO}, \ldots\} \).
  end.

It is necessary to generate a high variety of initial weight \( W_{\text{init}} \) sets in order to have a better solution. Otherwise \( W_{\text{temp}} \) would not change enough to avoid digging repetitively around the neighborhood of a specific local minima. The less variety of initial weights reduces the possibility to have a lower global minimum. This is why we generate initial weight sets randomly with a uniform distribution to get a high variety of initial weight sets.

### 3.2.2 Computation time complexity

The computation time complexity of PSO-NP is the total sum of the computation time complexity of steps 1, 2, and step 3.

The computation time complexity of step 1 is equal to that of SO. Under the SO optimization process, if we focus on determining one link weight, there are \( w_{\text{max}} \) combinations for each link weight. To determine \( L \) link weights, we do not decrement link weights and we only increment them at most \( Lw_{\text{max}} \) times to get the SO solution. Therefore, we have \( Lw_{\text{max}} \) weight-set combinations to determine \( L \) link weights for a given initial link weight set. For each weight set, we need to compute the network congestion ratio. Let \( O(X) \) be the computation time complexity to find the congestion ratio for each weight set. Moreover, we need to compute all the initial \( I_{\text{max}} \) sets. Therefore, the computation complexity of SO is \( O(Lw_{\text{max}}XI_{\text{max}}) \).

Consider the computation time complexity of steps 2 and 3. To determine the worst case congestion ratio we need to examine all possible link failure patterns.
3. PENALTY-AWARE MODEL FOR LINK WEIGHT OPTIMIZATION

including the no failure scenario, which gives us a total of \( L + 1 \) failure patterns. Therefore, the computation complexity to get the worst case congestion ratio is \( O(LX) \) for a link weight set in \( S \). We need this computation time complexity for every element of \( S \). Let \( S_{\text{max}} \) be the maximum number of elements in \( S \). The total computation time complexity of steps 2 and 3 is \( O(LXS_{\text{max}}) \).

In all, the PSO-NP computation time complexity is expressed by,

\[
O(Lw_{\text{max}}I_{\text{max}} + LXS_{\text{max}}).
\]  

(3.7)

\( O(X) \) can be evaluated by using the computation time complexity of Dijkstra’s algorithm, as the congestion ratio information can be deduced only after setting all the routing paths in the network. In Dijkstra’s algorithm, for each source node a shortest path tree is computed using a computation complexity of \( O(L + N \log N) \). Since Dijkstra’s algorithm is run for every node in the network, \( O(X) = O((L + N \log N) \times N) = O(NL + N^2 \log N) \).

By substituting \( X = NL + N^2 \log N \) into Eq. (3.7), the PSO-NP computation time complexity is expressed by,

\[
O((w_{\text{max}}I_{\text{max}} + S_{\text{max}})(NL^2 + N^2 L \log N)).
\]  

(3.8)

\( w_{\text{max}}, I_{\text{max}}, \) and \( S_{\text{max}} \) are the parameters that can be controlled. To enhance the optimization accuracy, we set \( w_{\text{max}}I_{\text{max}} > S_{\text{max}} \). In this case, the PSO-NP computation time complexity is expressed by \( O(w_{\text{max}}I_{\text{max}}(NL^2 + N^2 L \log N)) \), which is the same computation time complexity of SO.

3.2.3 Performance evaluation

The performances of the PSO-P, PSO-NP and SO are compared via simulations for both failure and non-failure scenarios through our heuristic approach. For each scheme we evaluate the congestion ratio under the worst case failure and non failure scenarios. Comparison metrics are respectively the reduction ratio of the worst congestion,

\[
\alpha_X = \frac{\max_{l \in F} r(W_{SO}, l) - \max_{l \in F} r(W_X, l)}{\max_{l \in F} r(W_{SO}, l)}.
\]  

(3.9)
3.2 PSO-NP: preventive start-time with no penalty

and the penalty under no failure,

\[ \beta_X = \frac{r(W_X,0) - r(W_{SO},0)}{r(W_{SO},0)}, \]  \hspace{1cm} (3.10)

where \( X = \text{PSO-P, PSO-NP and SO}. \)

Four sample networks, shown in Fig. 3.3 are used. The characteristics of the networks considered are shown in Table 3.1.

![Network 1](Network1.png) ![Network 2](Network2.png)
![Network 3](Network3.png) ![Network 4](Network4.png)

**Figure 3.3:** Examined network topologies.

**Table 3.1:** characteristics of networks

<table>
<thead>
<tr>
<th>Network type</th>
<th>No. of nodes</th>
<th>No. of links (bidirectional)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network 1</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>Network 2</td>
<td>12</td>
<td>22</td>
</tr>
<tr>
<td>Network 3</td>
<td>14</td>
<td>21</td>
</tr>
<tr>
<td>Network 4</td>
<td>10</td>
<td>20</td>
</tr>
</tbody>
</table>

Networks 1 and 2 mirror typical backbone networks \([17]\). Networks 3 is an Abilene network \([2]\). Network 4 is a random network generated via BRITE \([46]\).
3. PENALTY-AWARE MODEL FOR LINK WEIGHT OPTIMIZATION

For the given networks, link capacities are randomly generated with uniform distribution in the range of \((10U_C, 100U_C)\), where \(U_C\) [Gbit/s] is given a constant integer value. The traffic demands \(d_{pq}\) are also randomly generated with uniform distribution in the range of \((0, 100U_D)\), where \(U_D\) [Gbit/s] is a given constant integer value. \(U_D/U_C\) is determined so that we can get feasible solutions. \(z\) is set to 1 to evaluate the performance of a maximum number of weight sets. Tables 3.2 shows the results obtained for all sample networks. By definition both \(\alpha_{SO}\) and \(\beta_{SO}\) are equal to zero. Note that the congestion ratios are normalized by that of SO under no failure.

Table 3.2: Comparisons of PSO-NP and PSO-P performance from heuristic approach

<table>
<thead>
<tr>
<th>Network type</th>
<th>(r(0))</th>
<th>(\max_{i \in F} r(l))</th>
<th>(\alpha)</th>
<th>(\beta)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network 1</td>
<td>PSO-P</td>
<td>1.42</td>
<td>1.47</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>PSO-NP</td>
<td>1.00</td>
<td>1.50</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>SO</td>
<td>1.00</td>
<td>2.68</td>
<td>0.00</td>
</tr>
<tr>
<td>Network 2</td>
<td>PSO-P</td>
<td>1.25</td>
<td>1.56</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>PSO-NP</td>
<td>1.00</td>
<td>2.46</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>SO</td>
<td>1.00</td>
<td>7.78</td>
<td>0.00</td>
</tr>
<tr>
<td>Network 3</td>
<td>PSO-P</td>
<td>1.67</td>
<td>2.03</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>PSO-NP</td>
<td>1.00</td>
<td>2.37</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>SO</td>
<td>1.00</td>
<td>2.37</td>
<td>0.00</td>
</tr>
<tr>
<td>Network 4</td>
<td>PSO-P</td>
<td>1.85</td>
<td>1.83</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>PSO-NP</td>
<td>1.00</td>
<td>2.02</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>SO</td>
<td>1.00</td>
<td>5.41</td>
<td>0.00</td>
</tr>
</tbody>
</table>

In Table 3.2 for network 1 and network 4, we can state that PSO-NP reduces the worst case congestion ratio almost as much as PSO-P while suppressing the penalty which, PSO-P does not. PSO-P carries a penalty of 42% and 85% for networks 1 and 4 respectively when there is no failure. This shows the effectiveness of PSO-NP in reducing the worst congestion ratio while eliminating the penalty when there is no failure. For network 2, the reduction ratio of PSO-NP is lower than PSO-P’s but PSO-NP still has the advantage of keeping penalty to zero. For this case also, network operators can select one of either depending on the
3.2 PSO-NP: preventive start-time with no penalty

quality of service they need to deliver. Finally, for network 3, PSO-NP does not show any amelioration compared to SO. We study this case later in this paper.

For larger networks the effectiveness of PSO-NP depends on $I_{max}$ values. For larger values of $I_{max}$, $\alpha_{PSO-NP}$ may be boosted and even match $\alpha_{PSO-P}$ because more initial weight sets means more possible elements of $S$. We examined the performance of PSO-NP for the values of $I_{max}$ ranging from 1 to 10000 in our examined networks. The results are listed in Tables 3.3. As $I_{max}$ increases it is clear that PSO-NP performance gets better. This can be observed especially in network 1, 2 and 4.

<table>
<thead>
<tr>
<th>Network type</th>
<th>$I_{max}$</th>
<th>$\alpha_{PSO-P}$</th>
<th>$\alpha_{PSO-NP}$</th>
<th>$\beta_{PSO-P}$</th>
<th>$\beta_{PSO-NP}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network 1</td>
<td>1</td>
<td>0.18</td>
<td>0.09</td>
<td>0.28</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.27</td>
<td>0.12</td>
<td>0.27</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.44</td>
<td>0.32</td>
<td>0.16</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.45</td>
<td>0.44</td>
<td>0.41</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>10000</td>
<td>0.45</td>
<td>0.44</td>
<td>0.42</td>
<td>0.00</td>
</tr>
<tr>
<td>Network 2</td>
<td>1</td>
<td>0.62</td>
<td>0.00</td>
<td>0.43</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.62</td>
<td>0.00</td>
<td>0.43</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.62</td>
<td>0.00</td>
<td>0.43</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.58</td>
<td>0.00</td>
<td>0.47</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>10000</td>
<td>0.80</td>
<td>0.68</td>
<td>0.25</td>
<td>0.00</td>
</tr>
<tr>
<td>Network 3</td>
<td>1</td>
<td>0.003</td>
<td>0.00</td>
<td>0.72</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.003</td>
<td>0.00</td>
<td>0.72</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.003</td>
<td>0.00</td>
<td>0.62</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.003</td>
<td>0.00</td>
<td>0.75</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>10000</td>
<td>0.14</td>
<td>0.00</td>
<td>0.67</td>
<td>0.00</td>
</tr>
<tr>
<td>Network 4</td>
<td>1</td>
<td>0.37</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.37</td>
<td>0.00</td>
<td>0.43</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.63</td>
<td>0.60</td>
<td>0.85</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.69</td>
<td>0.60</td>
<td>0.85</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>10000</td>
<td>0.66</td>
<td>0.63</td>
<td>0.85</td>
<td>0.00</td>
</tr>
</tbody>
</table>
3. PENALTY-AWARE MODEL FOR LINK WEIGHT OPTIMIZATION

Table 3.4: Computation time comparison of SO, PSO-NP and PSO-P in seconds

<table>
<thead>
<tr>
<th>Network type</th>
<th>SO</th>
<th>PSO-NP</th>
<th>PSO-P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network 1</td>
<td>3157.54</td>
<td>3140.79</td>
<td>18026.84</td>
</tr>
<tr>
<td>Network 2</td>
<td>31603.00</td>
<td>32129.77</td>
<td>336610.75</td>
</tr>
<tr>
<td>Network 3</td>
<td>42136.04</td>
<td>43073.50</td>
<td>627295.30</td>
</tr>
<tr>
<td>Network 4</td>
<td>16291.27</td>
<td>16559.43</td>
<td>182786.63</td>
</tr>
</tbody>
</table>

The computation times of SO, PSO-NP, PSO-P are examined as shown in Table 3.4. The simulation is performed by using a Linux-based computer equipped with a 2.3 GHz Intel Core i5 Processor and 4 GB memory. For the examined networks PSO-NP computation time is close to that of SO as we expected, while PSO-P computation time is 2 to 10 times longer than that of PSO-NP.

PSO-NP is able to delete the penalty of PSO-P and in some cases show worst case congestion ratio reduction gain similar to PSO-P. However it does not guarantee a reduction of worst congestion ratio as its link weight set may sometimes be equal to that of SO. In the following section we introduce S-PSO in order to guarantee the worst congestion ratio reduction gain achieved by PSO-P while at the same time reducing the penalty of PSO-P under no failure case.

3.3 S-PSO: strengthened preventive start-time optimization

S-PSO finds a link weight set that minimizes the worst congestion ratio after failure as the PSO-P does, and at the same time reduces the penalty of the PSO-P under no failure. S-PSO is an enhanced version of PSO-P. In the optimization procedure of the conventional PSO to minimize \( \max_{l \in F} r(W, l) \) \([30]\), the weight set that firstly hits the optimal solution \( R(W_{PSO-P}) \) is kept as the solution.

S-PSO uses the same procedure as the conventional PSO \([30]\). However, in S-PSO, for any link weight set \( W^k_{S-PSO} \), where \( k \) is index, such that \( R(W^k_{S-PSO}) = R(W_{PSO-P}) \), we group \( W^k_{S-PSO} \) into a set of \( S \). For each \( W^k_{S-PSO} \in S \), we evaluate
3.3 S-PSO: strengthened preventive start-time optimization

Then, the best solution of the S-PSO weight set is selected by,

\[ W_{S-PSO} = \arg \min_{W_{S-PSO}^{k} \in S} r(W_{S-PSO}^{k}, 0). \]  (3.11)

As a consequence, we have the following relationships,

\[ R(W_{S-PSO}) = R(W_{PSO-P}) \]  (3.12)

and

\[ r(W_{S-PSO}, 0) \geq r(W_{PSO-P}, 0). \]  (3.13)

\( W_{S-PSO} \) reduces the penalty of the conventional PSO under no failure while minimising the worst congestion ratio after failure as \( W_{PSO-P} \) does. Similarly to PSO-NP S-PSO ILP is expressed as:

\[
\begin{align*}
\text{Objective} & \quad \min_{W} \{R(W) + \epsilon \cdot r(W, 0)\} \\
\text{s.t.} & \quad \sum_{j:(i,j) \in E_l} x_{ij}^{pq}(W,l) - \sum_{j:(i,j) \in E_l} x_{ji}^{pq}(W,l) = 1, \quad \forall p, q \in V, i = p, l \in F \\
& \quad \sum_{j:(i,j) \in E_l} x_{ij}^{pq}(W,l) - \sum_{j:(i,j) \in E_l} x_{ji}^{pq}(W,l) = 0, \quad \forall p, q \in V, i \neq p, q, l \in F \\
& \quad \sum_{j:(i,j) \in E_l} x_{ij}^{pq}(W,l) - \sum_{j:(i,j) \in E_l} x_{ji}^{pq}(W,l) = -1, \quad \forall p, q \in V, i = q, l \in F \\
& \quad \sum_{p,q \in V} d_{pq} x_{ij}^{pq}(W,0) \leq c_{ij}^0 \cdot r_{SO}, \quad \forall (i,j) \in E_0 \\
& \quad \sum_{p,q \in V} d_{pq} x_{ij}^{pq}(W,l) \leq c_{ij}^l \cdot r_{PSO}, \quad \forall (i,j) \in E_l, l \in F \\
& \quad 0 \leq f_{pq}^l(l) - x_{ij}^{pq}(W,l) \leq 1 - \delta_{ij}^q(l), \forall p, q \in V, (i, j) \in E_l, l \in F \\
& \quad x_{ij}^{pq}(W,l) \leq \delta_{ij}^q(l), \forall p, q \in V, (i, j) \in E_l, l \in F
\end{align*}
\]  (3.14a-h)
3. PENALTY-AWARE MODEL FOR LINK WEIGHT OPTIMIZATION

\[ 0 \leq \psi_{jq}(l) + w_{ij} - \psi_{iq}(l) \leq (1 - \delta_{ij}(l))U, \quad \forall q \in V, (i, j) \in E, l \in F \]  \hspace{1cm} (3.14i)

\[ 1 - \delta_{ij}(l) \leq \psi_{jq}(l) + w_{ij} \leq \psi_{iq}(l), \quad \forall q \in V, (i, j) \in E, l \in F \]  \hspace{1cm} (3.14j)

\[ f_{ij}^q(l) \geq 0, \forall p, q, i \in V, l \in F \]  \hspace{1cm} (3.14k)

\[ \delta_{ij}^q(l) \in \{0, 1\}, \forall q \in V, (i, j) \in E, l \in F \]  \hspace{1cm} (3.14l)

\[ 1 \leq w_{ij} \leq w_{\text{max}}, \forall (i, j) \in E, l \in F. \]  \hspace{1cm} (3.14m)

The Heuristic formulation can be easily deduced from that of PSO-NP and is described in the following procedure.

### 3.3.1 Heuristic procedure

Analogically to the heuristic procedure of PSO-NP 3.2.1, S-PSO’s procedure to calculate \( W_{S-PSO} \) is presented and divided into the following three steps.

- **Step 1:** Get \( W_{PSO-P} \) and collect all weight sets that generate the same \( \max_{l \in F} r(W, l) \). These weight sets are included into a set of \( S' \).

- **Step 2:** For each \( W_{S-PSO}^k \in S' \), evaluate \( r(W_{S-PSO}^k, 0) \).

- **Step 3:** Return \( W_{S-PSO} \) that minimizes \( r(W_{S-PSO}^k, 0) \) using Eq. (3.11).

The determination of the suitable link weight set is performed as explained in PSO-NP heuristic procedure.

### 3.3.2 Performance evaluation

The performances of SO, S-PSO, and PSO-P are compared via simulations for both failure and non-failure scenarios. For each scheme, we evaluate \( \max_{l \in F} r(l) \) and \( r(0) \).

Traffic demands between sources and destinations are set randomly from 0 to 100 units. Link capacities are set randomly from 150 to 200 units. \( w_{\text{max}} \) is set to 1000. Performance evaluation is made for network 1 and 2.
3.4 Summary

Table 3.5 shows the worst congestion ratio under failure, $\max_{l \in F} r(l)$, and the congestion ratio under no failure, $r(0)$. As we expected, S-PSO gives the same results as the conventional PSO under the worst case failure. For the no failure case S-PSO shows a lower congestion ratio than the conventional PSO. The reduction effect of the congestion ratio under no failure of S-PSO compared to that of the conventional PSO is defined by,

$$\delta = \frac{r_{PSO-P}(0) - r_{S-PSO}(0)}{r_{PSO-P}(0)}.$$  \hspace{1cm} (3.15)

We obtain $\delta = 0.15$ and $0.29$ for networks 1 and 2, respectively.

Table 3.5: Comparisons of three schemes under failure and no failure scenarios

<table>
<thead>
<tr>
<th>Network type</th>
<th>$r(0)$</th>
<th>$\max_{l \in F} r(l)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SO</td>
<td>1.00</td>
<td>1.47</td>
</tr>
<tr>
<td>Network 1</td>
<td>S-PSO</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>PSO-P</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.44</td>
</tr>
<tr>
<td>SO</td>
<td>1.00</td>
<td>1.85</td>
</tr>
<tr>
<td>Network 2</td>
<td>S-PSO</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>PSO-P</td>
<td>1.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.70</td>
</tr>
</tbody>
</table>

3.4 Summary

In order to reduce the penalty $\beta$ and increase the worst congestion reduction ratio $\alpha$, we presented PSO-NP and S-PSO. In the considered sample networks, PSO-NP and S-PSO are able to outperform PSO-P with lower values $\beta$. They can also give values $\alpha$ closed to PSO-P’s. However, PSO-NP and S-PSO cannot guarantee better sets as their solutions are subsets of those of SO and PSO respectively. Furthermore, PSO-NP only regroups weight sets that keep $r(0) = r_{SO}(0) = 1.0$ whereas S-PSO regroups those that keep $\max_{l \in F} r(l) = R(W_{PSO})$. Sub-optimal sets that are suitable under both failure and non failure scenarios are dropped. For example a set with $r(0) = 1.001$ and $\max_{l \in F} r(l) = R(W_{PSO}) + 0.001$ will be dropped by both PSO-NP and PSO-P. Under a situation where PSO-NP and S-PSO share the same solution with SO and PSO-P respectively, the above sub-optimal weight sets can present better applicability. Therefore a scheme that
3. PENALTY-AWARE MODEL FOR LINK WEIGHT OPTIMIZATION

considers sub-optimal sets by balancing $r(0)$ (or $\beta$) and $\max_{l \in F} r(l)$ (or $\alpha$) needs to be investigated.
Chapter 4

Generalization of penalty-aware model of link weight optimization

4.1 Overview

In this chapter we generalize the penalty-aware model of link weight optimization by presenting a general preventive start-time optimization (GPSO). GPSO balances the worst congestion ratio reduction gain and the penalty under failure in order to find a more flexible sub-optimal weight sets. In other words GPSO is designed to give network operators the latitude to get a suitable link weight set that considers the balance between $r(0)$ and $\max_{l \in F} r(l)$ instead of focusing on one of them as SO, PSO-P, PSO-NP and S-PSO does. GPSO ILP formulation derives from Eq. (4.1a)-(4.1m) and can be expressed as:

Objective \( \min_W \{ (1 - \gamma) \cdot r(W,0) + \gamma \cdot R(W) \} \), \hspace{1cm} (4.1a)

s.t. \[ \sum_{j:(i,j) \in E_i} x_{ij}^{pq}(W,l) - \sum_{j:(i,j) \in E_i} x_{ji}^{pq}(W,l) = 1, \]
\[ \forall p, q \in V, i = p, l \in F \] \hspace{1cm} (4.1b)

\[ \sum_{j:(i,j) \in E_i} x_{ij}^{pq}(W,l) - \sum_{j:(i,j) \in E_i} x_{ji}^{pq}(W,l) = 0, \]
\[ \forall p, q \in V, i \neq p, q, l \in F \] \hspace{1cm} (4.1c)
4. GENERALIZATION OF PENALTY-AWARE MODEL OF LINK WEIGHT OPTIMIZATION

\[
\sum_{j: (i,j) \in E_l} x_{ij}^{pq}(W,l) - \sum_{j: (i,j) \in E_i} x_{ji}^{pq}(W,l) = -1, \\
\forall p, q \in V, i = q, l \in F
\]

(4.1d)

\[
\sum_{p,q \in V} d_{pq} x_{ij}^{pq}(W,0) \leq c_{ij}^0 \cdot r_{SO}, \\
\forall (i,j) \in E_0
\]

(4.1e)

\[
\sum_{p,q \in V} d_{pq} x_{ij}^{pq}(W,l) \leq c_{ij}^l \cdot r_{PSO}, \\
\forall (i,j) \in E_l, l \in F
\]

(4.1f)

\[
x_{ij}^{pq}(W,l) \leq \delta_{ij}^q(l), \forall p, q \in V, (i,j) \in E_l, l \in F
\]

(4.1g)

\[
0 \leq f_{pq}^i(l) - x_{ij}^{pq}(W,l) \leq 1 - \delta_{ij}^q(l), \forall p, q \in V, (i,j) \in E_l, l \in F
\]

(4.1h)

\[
0 \leq \psi^{jq}(l) + w_{ij} - \psi^{iq}(l) \leq (1 - \delta_{ij}^q(l))U, \\
\forall q \in V, (i,j) \in E_l, l \in F
\]

(4.1i)

\[
1 - \delta_{ij}^q(l) \leq \psi^{jq}(l) + w_{ij} \leq \psi^{iq}(l), \\
\forall q \in V, (i,j) \in E_l, l \in F
\]

(4.1j)

\[
f_{pq}^i(l) \geq 0, \forall p, q, i \in V, l \in F
\]

(4.1k)

\[
\delta_{ij}^q(l) \in \{0, 1\}, \forall q \in V, (i,j) \in E_l, l \in F
\]

(4.1l)

\[
1 \leq w_{ij} \leq w_{\text{max}}, \forall (i,j) \in E_l, l \in F.
\]

(4.1m)

where \(0 \leq \gamma \leq 1\). \(\gamma\) is a parameter defined by network operators.

With the appropriate \(\gamma\), GPSO recovers SO, PSO-NP, S-PSO and PSO-P. For \(\gamma = 0\), GPSO gives the same result as SO because only the congestion ratio under no failure is considered. When \(\gamma\) is set to 1, GPSO minimizes the worst case congestion ratio which PSO-P does. In this case the weight that has the lowest \(r_{SO}(0)\) among the ones that minimizes the worst congestion ratio is chosen as in S-PSO [32]. Finally, for values of \(\gamma\) set between 0 to 1, network operators can find the best couple \((r(0), \max_{l \in F} r(l))\) instead of just targeting one of them as in SO and PSO-P.
4.2 Heuristic procedure

In total GPSO is summarised as follows:

\[
\text{GPSO} = \begin{cases} 
\text{SO}, & \text{if } \gamma = 0 \\
\text{PSO-NP}, & \text{if } 0 < \gamma \ll \frac{r_{SO}}{r_{PSO}} \\
\text{PSO-LP}, & \text{if } \frac{r_{SO}}{r_{PSO}} < \gamma < 1 \\
\text{PSO-P}, & \text{if } \gamma = 1.
\end{cases}
\]

Where PSO-LP (PSO with limited penalty) determines sub-optimal link weight sets for congestion ratio reduction under both failure non failure scenarios. That is the weight set with the best couple \((r(0), \max_{l \in F} r(l))\) from a network operator point of view.

Unfortunately, as mentioned previously ILP is difficult to resolve for large topologies. A heuristic formulation of GPSO is therefore required.

4.2 Heuristic procedure

In our heuristic algorithm we directly evaluate the balance between \(\alpha\) and \(\beta\) because finding the weight set that gives the best \((r(0), \max_{l \in F} r(l))\) is equivalent to finding the one with the best \((\alpha, \beta)\). For that, we allow the congestion ratio under no failure \(r(0)\) to have a range. That range is set as:

\[
r_{SO}(0) \leq r(0) \leq r_{SO}(0) + \delta, \quad \delta > 0,
\]

where \(\delta\) is given. In other words, under no failure the penalty is not kept at zero like PSO-NP but bounded as:

\[
0 \leq \beta \leq \beta_{\text{upp}},
\]

where

\[
\beta_{\text{upp}} = \frac{\delta}{r_{SO}(0)}
\]

is the upper bound penalty.

The aim of this section is to show the relationship between \(\alpha\) and \(\beta_{\text{upp}}\) in order to help the network operator choose a suitable couple \((\alpha, \beta_{\text{upp}})\) and its corresponding weight set. The upper bound penalty is the maximal allowable penalty when there is no failure in the network. For a given value of \(\beta_{\text{upp}}\), GPSO
finds the weight set that reduces most $\max_{l \in F} r(l)$ boosting $\alpha$ as a consequence. $\beta_{PSO-P}$ is the highest possible penalty to pay under a no failure scenario compared to SO scheme. Therefore the range of $\beta_{upp}$ is:

$$0 \leq \beta_{upp} \leq \beta_{PSO-P},$$

(4.5)

Since $\beta_{upp}$ represents the maximal allowable penalty under no failure in the network, the solution weight set of GPSO can be expressed as:

$$W_{GPSO_{\beta_{upp}}} = \arg \max_{W} \alpha(W).$$

(4.6)

For values of $\beta_{upp}$ moving from 0 to $\beta_{PSO-P}$ we obtain respectively PSO-NP, PSO-LP and PSO-P schemes. In general, we have:

$$\text{GPSO} = \begin{cases} 
\text{PSO-NP,} & \text{if } \beta_{upp} = 0 \\
\text{PSO-LP,} & \text{if } 0 < \beta_{upp} < 1, \beta_{upp} < \beta_{PSO-P} \\
\text{S-PSO,} & \text{if } \beta_{upp} = \beta_{PSO-P}.
\end{cases}$$

4.3 Performance evaluation

We examined the performance of GPSO in network 3 by drawing the graph that gives the relationship between $\alpha$ and $\beta_{upp}$. This graph is shown in Fig. 4.1 and for $\beta_{upp} = 0.0015$ we can achieve a reduction ratio equal to that of PSO-P. This demonstrate that Fig. 4.1 can help network operators find weight sets with marginal or negligible penalty with maximal reduction of the worst congestion ratio. In addition, for values of $\beta_{upp}$ equal to 0 and $\beta_{PSO-P}$, GPSO matches PSO-NP and PSO-P reduction ratio, respectively. GPSO therefore generalizes the previous schemes and goes further to provide a more flexible weight set depending on the upper bound penalty when there is no failure.

4.4 Summary

GPSO determines a weight set that keeps the penalty negligible while substantially reducing the worst case congestion ratio as it tunes low values of the penalty to find the weight set that minimizes the worst case congestion ratio. GPSO also clarifies the relationship between the manageable congestion ratio and the penalty.
under no failure as well as the worst congestion reduction ratio. This gives to network operators the latitude to determine which manageable congestion ratio should be to set according to their congestion requirements under failure and non failure cases.

However, GPSO still contains the penalty. As a result it may not meet low and strict manageable congestion conditions. In the second part of this thesis we propose a link-duplication (LD) model that aims to suppress link failure in the first place in order to always meet the manageable congestion under single-link failure. For this purpose we consider the duplication or reinforcement of links which is broadly used to make networks reliable.
4. GENERALIZATION OF PENALTY-AWARE MODEL OF LINK WEIGHT OPTIMIZATION
Chapter 5

Link-duplication model for link weight optimization

5.1 Overview

This chapter introduces a link-duplication model where link weights are optimized in order to reduce the number of protected links with the condition of always maintaining the manageable congestion under single-link failure scenario.

Link duplication methods such as 1+1 or 1:1 protection are widely used in networks for reliability. Link duplication provides fast recovery as only switching from the failed link to the backup link will hide the failure at upper layers. However, due to CAPEX constraints, every link cannot be duplicated. Giving priority to some selected links makes sense. As mentioned above, traffic routes are determined by link weights that are configured in advance. Therefore, choosing an appropriate set of link weights may reduce the number of links that actually need to be duplicated in order to keep a manageable congestion under failure and also non failure cases. Now, PSO also determines the link failure which creates the worst congestion after failure. Since by duplicating this link we can assume it no more fails, PSO can be used to find the smallest number of links to protect so as to guarantee a manageable congestion under any single-link failure. In this chapter we show how to find the suitable links to duplicate through a link weight optimization scheme when the manageable congestion ratio is given.
5. LINK-DUPLICATION MODEL FOR LINK WEIGHT OPTIMIZATION

5.2 Link weight optimization scheme with link duplication [31]

PSO determines the weight set that minimizes the worst congestion under failure. PSO enable us to also find the critical link that creates the worst congestion ratio in case it fails. In S-PSO and PSO-NP, PSO algorithm [30] is applied to the whole set of links in the network. In fact the same applies to a specific set of links in the network. PSO can be reformulated as follow: For any given failure pattern \( H \subseteq F \), PSO identifies the worst-case failure \( l_H \in H \) and determines a weight set that minimizes the worst-case congestion ratio in case that \( l_H \) fails. In other words, for a given \( H \), PSO provides \( l_H \) and

\[
W_{PSO_H} = \arg \min_W R_H(W). \tag{5.1}
\]

Let us denote \( r(W_{PSO_H}, l_H) \) as \( r_{l_H} \). When \( W_{PSO_H} \) is applied to the network, any single-link failure \( l \in H \) generates a congestion ratio at most equal to \( r_{l_H} \) by definition of PSO. Therefore,

\[
r_{l_H} \geq r(W_{PSO_H}, l) \quad \forall l \in H. \tag{5.2}
\]

Now, if we apply PSO only on the singleton \( \{l_H\} \) which represents the worst-case failure in \( H \), any solution weight set \( W_{PSO_{\{l_H\}}} \) will satisfy the same minimal congestion ratio \( r_{l_H} \) since for a given failure the optimized congestion ratio is unique. As a result we have,

\[
r(W_{PSO_{\{l_H\}}}, l_H) = r_{l_H}. \tag{5.3}
\]

Therefore, for any weight set \( W \) we always have,

\[
r(W, l_H) \geq r_{l_H}. \tag{5.4}
\]

5.2.1 Formulation

In this section we present and formulate a link-duplication optimization (LDO) scheme. LDO uses PSO to determine the minimal number of links that need to be reinforced so as to keep a manageable congestion ratio under any single-link
5.2 Link weight optimization scheme with link duplication [31]

failure scenario. Network operators can set the upper bound of the congestion ratio that they can allow under any single-link failure scenario including the non-failure case. Let us denote this upper bound congestion ratio as $r_{\text{upp}}$.

$r_{\text{upp}}$ would be set by considering user’s QoE (Quality of Experience) by network operators. For premium and expensive services which require a high QoE, $r_{\text{upp}}$ would be set at a low value while for cheaper services it would be set higher. However, it is difficult to exactly determine the QoE of a specific service since it varies with the environment which may not be stable. This is the case in mobile communication [25]. Moreover, it is still challenging to build a unified model to measure and evaluate the QoE of different services and also meet requirements that may highly vary between users [61].

Since $r_{SO}(0)$ is the minimal possible congestion ratio under no failure, $r_{\text{upp}}$ can be expressed by,

$$r_{\text{upp}} = (1 + \gamma)r_{SO}(0), \quad \gamma \geq 0,$$  \hspace{1cm} (5.5)

where $\gamma$ represents a margin defined by the network operator. $r_{\text{upp}}$ is $\gamma r_{SO}(0)$ higher than the minimal congestion ratio of $r_{SO}(0)$.

When $\gamma$ is given, we can find a set of failure indices $F_{\text{sol}}$ on which PSO is able to find $W_{PSO,F_{\text{sol}}}$ such that

$$r_{SO}(0) \leq r_{l_{F_{\text{sol}}}} \leq r_{\text{upp}}.$$  \hspace{1cm} (5.6)

$F_{\text{sol}}$ and $W_{PSO,F_{\text{sol}}}$ do exist since applying PSO on the singleton $H = \{0\}$ is equivalent to applying SO in the network. In Eq. (5.1), we replace $H$ by $\{0\}$ and we have $R_{\{0\}}(W) = r(W,0)$ which means that from Eq. (5.1) we get,

$$W_{PSO_{\{0\}}} = \arg \min_{W} R_{\{0\}}(W) = \arg \min_{W} r(W,0) = W_{SO}.$$  \hspace{1cm} (5.7)

As a consequence, $F_{\text{sol}}$ exists and contains at least the singleton $\{0\} \subset F$.

As long as every link $l \in F \setminus F_{\text{sol}}$ is reinforced, PSO can determine a weight set $W_{PSO,F_{\text{sol}}}$ at the start time such that under any single-link failure $l \in F_{\text{sol}}$ the network congestion ratio is kept lower or equal to $r_{\text{upp}}$, which is manageable.

When $r_{\text{upp}}$ is given, $F_{\text{sol}}$ is constructed in Algorithm 1. $X$ represents the set of links to protect or reinforce so as to keep the manageable congestion ratio $r_{\text{upp}}$.  

45
5. LINK-DUPLICATION MODEL FOR LINK WEIGHT OPTIMIZATION

**Algorithm 1** Determination of $F_{\text{sol}}$ and $W_{\text{PSO}_F_{\text{sol}}}$

We run PSO on $F_0 = F$ to determine a weight set $W_{\text{PSO}_F_0}$ that minimizes $R_{F_0}$ with the worst-case failure pattern being $l_{F_0}$. We have $r(W_{\text{PSO}_F_0}, l_{F_0}) = r_{l_{F_0}}$. If $r_{l_{F_0}} \leq r_{\text{upp}}$, the network operator can deal with the worst-case failure. In this case, $F_{\text{sol}} = F_0$, $W_{\text{PSO}_F_{\text{sol}}} = W_{\text{PSO}_F_0}$, and the number of links to reinforce is zero.

Now, if $r_{l_{F_0}} > r_{\text{upp}}$, we put $l_{F_0}$ in a set $X$. $X$ represents the set of links to reinforce. We then run PSO on $F_1 = F \setminus X$ to determine $W_{\text{PSO}_F_1}$ that minimizes $R_{F_1}$ with the worst-case failure pattern being $l_{F_1}$. We have $r(W_{\text{PSO}_F_1}, l_{F_1}) = r_{l_{F_1}}$. We can say that $r_{l_{F_0}} \geq r_{l_{F_1}}$ due to $F_1 \subseteq F_0$. If $r_{l_{F_1}} \leq r_{\text{upp}}$ then the network operator can deal with all failures in $F_1$. In this case $F_{\text{sol}} = F_1$ and $W_{\text{PSO}_F_{\text{sol}}} = W_{\text{PSO}_F_1}$.

We repeat this process by constructing $F_m$ and we determine $W_{\text{PSO}_F_m}$ until $r_{l_{F_m}} \leq r_{\text{upp}}$ with $r_{l_{F_{m-1}}} > r_{\text{upp}}$. This convergence is assured by the fact that, if PSO is applied on the singleton $\{0\}$, it will yield $r_{SO}(0)$ which is lower or equal to $r_{\text{upp}}$.

At the end of the procedure we have,

$$r_{l_{F_0}} \geq r_{l_{F_1}} \geq r_{l_{F_2}} \ldots \geq r_{l_{F_{m-1}}} > r_{l_{F_m}},$$

(5.8)

and

$$r_{SO}(0) \leq r_{l_{F_m}} \leq r_{\text{upp}} < r_{l_{F_{m-1}}}.$$

(5.9)

The final $X$ is,

$$X = \{l_{F_0}, l_{F_1}, \ldots, l_{F_{m-1}}\},$$

(5.10)

and we have $F_{\text{sol}} = F_m$, $F \setminus F_{\text{sol}} = X$, and $W_{\text{PSO}_F_{\text{sol}}} = W_{\text{PSO}_F_m}$.

**Theorem 1** For a given $r_{\text{upp}}$, let $F_{\text{sol}}$ be the set returned by Algorithm 1 and $X = F \setminus F_{\text{sol}}$. $|X|$ is minimal.

**Proof:** Let us show that $|X|$ is minimal by absurd. Suppose that there is another set $X_1$ and a weight set $W_{X_1}$ such that:

$$|X_1| < |X| \text{ and } r_{SO}(0) \leq r(W_{X_1}, t) \leq r_{\text{upp}} \forall t \in F \setminus X_1.$$  

(5.11)

Since $X_1 \neq X$ there is $l_{F_x} \in X \setminus X_1$, $0 \leq x \leq m - 1$. $X \setminus X_1 \subseteq F \setminus X_1$ means
5.2 Link weight optimization scheme with link duplication

\( l_{F_x} \in F \setminus X_1 \). Therefore, from Eq. (5.11),

\[
\begin{align*}
    r_{SO}(0) & \leq r(W_{X_1}, l_{F_x}) \leq r_{upp}.  \\
    (5.12)
\end{align*}
\]

If we apply PSO only on \( \{l_{F_x}\} \), from Eq. (5.4), for any weight set \( W, r(W, l_{F_x}) \geq r_{l_{F_x}} \). For \( W_{X_1} \) in particular, \( r(W_{X_1}, l_{F_x}) \geq r_{l_{F_x}} \). Since \( l_{F_x} \in X \), from Eq. (5.8), \( r_{l_{F_x}} \geq r_{l_{F_{x-1}}} > r_{upp} \). Therefore \( r(W_{X_1}, l_{F_x}) \geq r_{l_{F_x}} > r_{upp} \) which contradicts the fact \( r(W_{X_1}, l_{F_x}) \leq r_{upp} \).

Since \( X \) depends on \( r_{upp} \) and \( r_{upp} \) is determined by \( \gamma \), \( X \) is a function of \( \gamma \) and is denoted by \( X(\gamma) \).

5.2.2 Performance evaluation

The performances of LDO is examined. The measure variable is the protection ratio, \( P(\gamma) \), which is defined as the ratio of the number of links to protect to the total number of links,

\[
    P(\gamma) = \frac{|X(\gamma)|}{|E|},
\]

where \( \gamma \) is given.

Obviously a low value of \( P(\gamma) \) is desirable as we want to minimize \( |X(\gamma)| \). PSO minimizes the worst possible congestion ratio, when any single failure in \( F(= F_0) \) is considered. The minimal value is \( r_{l_{F_0}} \), where \( l_{F_0} \) is the link failure that generates \( r_{l_{F_0}} \). We consider \( r_{upp} \leq r_{l_{F_0}} \) because otherwise, \( r_{l_{F_0}} \) would be manageable which is trivial as any link duplication is not needed. The range of \( r_{upp} \) is defined by:

\[
    r_{SO}(0) \leq r_{upp} \leq r_{l_{F_0}}.
\]

As a result, the range that should be considered for \( \gamma \) is,

\[
    0 \leq \gamma \leq \gamma_{max},
\]

where,

\[
    \gamma_{max} = \frac{r_{l_{F_0}}}{r_{SO}(0)} - 1.
\]
5. LINK-DUPLICATION MODEL FOR LINK WEIGHT OPTIMIZATION

Evaluation of LDO is made by using four sample networks as shown in Fig. 3.3. Traffic demands between sources and destinations are also set randomly as in S-PSO. The protection ratios and $\gamma_{\text{max}}$ for each examined network are shown in Table 5.1. Table 5.1 observes that the protection ratio decreases when $\theta$ increases.

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>Network 1</th>
<th>Network 2</th>
<th>Network 3</th>
<th>Network 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.91</td>
<td>1.00</td>
<td>0.95</td>
<td>0.50</td>
</tr>
<tr>
<td>0.01</td>
<td>0.82</td>
<td>1.00</td>
<td>0.95</td>
<td>0.50</td>
</tr>
<tr>
<td>0.02</td>
<td>0.82</td>
<td>1.00</td>
<td>0.95</td>
<td>0.40</td>
</tr>
<tr>
<td>0.05</td>
<td>0.82</td>
<td>1.00</td>
<td>0.95</td>
<td>0.40</td>
</tr>
<tr>
<td>0.10</td>
<td>0.82</td>
<td>1.00</td>
<td>0.91</td>
<td>0.30</td>
</tr>
<tr>
<td>0.20</td>
<td>0.72</td>
<td>0.95</td>
<td>0.67</td>
<td>0.30</td>
</tr>
<tr>
<td>0.40</td>
<td>0.72</td>
<td>0.82</td>
<td>0.57</td>
<td>0.15</td>
</tr>
<tr>
<td>0.50</td>
<td>0.54</td>
<td>0.64</td>
<td>0.57</td>
<td>0.10</td>
</tr>
<tr>
<td>0.60</td>
<td>0.37</td>
<td>0.45</td>
<td>0.43</td>
<td>0.10</td>
</tr>
<tr>
<td>0.80</td>
<td>0.18</td>
<td>0.05</td>
<td>0.33</td>
<td>0.05</td>
</tr>
<tr>
<td>0.90</td>
<td>0.18</td>
<td>0.05</td>
<td>0.14</td>
<td>0.05</td>
</tr>
<tr>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\gamma_{\text{max}}$</td>
<td>0.41</td>
<td>0.63</td>
<td>0.48</td>
<td>0.75</td>
</tr>
</tbody>
</table>

where $\gamma = \theta \gamma_{\text{max}}$. Since $r_{\text{upp}} = (1 + \gamma)r_{SO}(0)$, $r_{\text{upp}}$ is $\theta \gamma_{\text{max}} r_{SO}(0)$ higher than $r_{SO}(0)$. In other terms, $r_{\text{upp}}$ is $\theta \gamma_{\text{max}} \times 100$ percent higher than $r_{SO}(0)$. In the following of this section, we state our observations for some selected results in Table 5.1 to help readers to interpret our results.

In network 4, for $\theta = 0$ the protection ratio is 0.5. As $\theta = 0$ means $r_{\text{upp}} = r_{SO}(0)$, we can state that the minimal congestion ratio can be kept under any single-link failure if we reinforce 50% of all links. From Table 3.1, network 4 has 20 links. Therefore, under any single-link failure, $r_{SO}(0)$ can be kept by only reinforcing 10 ($=20 \times 0.5$) links. For $\theta = 0.1$, the protection ratio is 0.30. As $\gamma_{\text{max}} = 0.75$, which is indicated at the bottom line of Table 5.1, $r_{\text{upp}}$ is therefore 7.5% ($=\theta \gamma_{\text{max}} \times 100$) higher than $r_{SO}(0)$. In this case reinforcing only 6 ($=20 \times 0.3$) links is enough to maintain a congestion ratio 7.5% higher than $r_{SO}(0)$ under any single-link failure scenario.
In network 1, for $\theta = 0.9$, the protection ratio is 0.18. As $\gamma_{\text{max}} = 0.41$, $r_{\text{upp}}$ is 37% ($=\theta \gamma_{\text{max}} \times 100$) higher than $r_{SO}(0)$. Moreover from Table 3.1 network 1 has 11 links. Therefore reinforcing 18% of all links (2 links) is enough to handle any single-link failure scenario as long as manageable congestion ratio is 37% higher than $r_{SO}$.

In network 2, when $\theta = 0.8$ the protection ratio is 0.05. Since $\gamma_{\text{max}} = 0.63$, $r_{\text{upp}}$ is 51% ($=\theta \gamma_{\text{max}} \times 100$) higher than $r_{SO}(0)$. From Table 3.1 network 2 has 22 links. Therefore, if a congestion ratio 51% higher than $r_{SO}(0)$ is allowed we can tackle any single-link failure case by only reinforcing 5% of all links (1 link).

Finally, in network 3, when $\theta = 0.9$ the protection ratio is 0.14. Since $\gamma_{\text{max}} = 0.48$, $r_{\text{upp}}$ is 39% ($=\theta \gamma_{\text{max}} \times 100$) higher than $r_{SO}(0)$. From Table 3.1 network 3 has 21 links. Therefore, if a congestion ratio 39% higher than $r_{SO}(0)$ is allowed we can tackle any single-link failure case by only reinforcing 14% of all links (3 links).

Overall, LDO reduces the number of links to reinforce when the manageable congestion ratio is given. It also provides a suitable weight set that will keep the network congestion ratio lower or equal to the manageable value under any single-link failure scenario.

However, link weights are fixed before determining links to duplicate. As the selection of links for protection depends on the fixed link weights, some suitable protection patterns, which are not considered with other possible link weights, might be skipped leading to overprotection. In the next section we present a scheme that considers multiple protection scenarios before optimizing link weights in order to further reduce the overall number of protected links.

### 5.3 Enhanced link weight optimization scheme with link duplication

We present an enhanced version of LDO, which we call E-LDO. E-LDO considers protection scenarios before optimizing link weights. For each protection candidate, we use PSO to calculate a link weight set that minimizes the worst or highest congestion ratio for any link failure related to non-protected links. The
5. LINK-DUPLICATION MODEL FOR LINK WEIGHT OPTIMIZATION

protection scenario that lowers the manageable congestion ratio with the lowest number of protected links is considered as the solution.

Let us consider as an example the network shown in Fig. 5.1. First, we run PSO on the network when there is no link reinforcement to minimize the worst congestion ratio. We find the link weight set \( W_0 = \{w_{10}, w_{20}, w_{30}, w_{40}\} \) that minimizes the worst congestion ratio that is denoted as \( r_0 \). Then we run PSO on the network where only \( l_1 \) is reinforced to minimize the worst congestion and we obtain the couple \( (W_1, r_1) \), where \( W_1 = \{w_{11}, w_{21}, w_{31}, w_{41}\} \). We repeat this process for \( l_2, l_3 \) and \( l_4 \) to obtain respectively the couples \( (W_2, r_2) \), \( (W_3, r_3) \) and \( (W_4, r_4) \). Let

\[
\text{r}_{\text{max}} = \min_{0 \leq i \leq 4} r_i, \tag{5.17}
\]

and \( l_{\text{max}} \) be the link protection that generates \( r_{\text{max}} \). If

\[
\text{r}_{\text{max}} \leq r_{\text{upp}}, \tag{5.18}
\]

by protecting only \( l_{\text{max}} \) and applying \( W_{\text{max}} \), the manageable congestion is guaranteed even under failure and non failure scenarios. Otherwise we conclude that a single protection is not enough. In that case we fix \( l_{\text{max}} \) as protected and we
5.3 Enhanced link weight optimization scheme with link duplication

5.3.1 Formulation

At first we apply PSO on \( F \) to check if \( r_{\text{upp}} \) can be satisfied without any link duplication. Otherwise, at least one link duplication is required. In that case, let \( Y = \{j_0, j_1, \ldots, j_L\} \) be the set of protection indices, where, for each element \( j \in Y \), \( l_j \) is reinforced. We denote \( X_2 \) as the set of reinforced link. \( X_2 \) is initialy set as an empty set. For every provisional reinforcement of link \( l_j \), \( j \in Y \), apply PSO on \( Y_{l_j} = Y \setminus \{l_j\} \), to minimize the worst congestion ratio and get \( r_{l_j} \). Let \( W_{\text{PSO}} \) represent the corresponding weight set under which \( r_{l_j} \) is obtained. \( r_{l_j} \) is the minimized worst congestion ratio when \( l_j \) is duplicated. From a link setting point of view the best protection scenario \( l_{\text{min}} \) is a link such that,

\[
l_{\text{min}} = \arg \min_{l_j \in Y} r_{l_j}. \tag{5.19}
\]

We set \( l_{\text{min}} \) as protected and we denote it as \( l_1 \) for simplicity. If \( r_{l_1} \leq r_{\text{upp}} \), only protecting \( l_1 \) is enough to keep the manageable congestion ratio. Otherwise, we put \( l_1 \) in \( X_2 \). Then we replace \( Y \) by \( Y \setminus X_2 \).

Then, we repeat the same procedure above by testing all provisional reinforced links in \( Y \setminus X_2 \) and update \( X_2 \) if the newly found \( r_{l_2} > r_{\text{upp}} \). We can also say that \( r_{l_1} \geq r_{l_2} \) due to \( Y \setminus X_2 \subset Y \). \( r_{\text{upp}} \) will be reached as protecting all the links gives the minimal congestion ratio, \( r_{SO}(0) \), which is lower than \( r_{\text{upp}} \). Let \( l_m \) represent the last protected link before we reach \( r_{\text{upp}} \). At the end of the procedure we have,

\[
r_{l_1} \geq r_{l_2} \geq r_{l_3} \cdots \geq r_{l_{m-1}} > r_{l_m}, \tag{5.20}
\]

and

\[
r_{SO}(0) \leq r_{l_m} \leq r_{\text{upp}} < r_{m-1}. \tag{5.21}
\]
5. LINK-DUPLICATION MODEL FOR LINK WEIGHT OPTIMIZATION

The resulting $X_2$ is,

$$X_2 = \{l_1, l_2, \cdots, l_{m-1}\}. \tag{5.22}$$

As $r_{\text{upp}}$ is a function of $\gamma$, $X_1$ in Eq. (5.10) and $X_2$ in Eq. (5.22) are both functions of $\gamma$ and are denoted as $X_1(\gamma)$, $X_2(\gamma)$ respectively.

5.3.2 Performance evaluation

The performances of LDO and E-LDO are compared. We also include an exhaustive search scheme that considers all possible combinations of link protections to minimize the worst congestion ratio when the number of protected links is given. The number of all possible combinations for the $n$-link protection scenario is $L_n$. This scheme is not practical for large networks as the computation time grows exponentially with the number of links in the network. We use the exhaustive search scheme as a reference scheme to check the performance of LDO and E-LDO for a small number of link protections.

A value to be evaluated is the protection ratio, $P(i, \gamma)$, which is defined as the ratio of the number of links protections to the total number of links. It is expressed as:

$$P(i, \gamma) = \frac{|X_i(\gamma)|}{|E|}, \tag{5.23}$$

where $\gamma$ is given, and $i$ represents the scheme used. Obviously a low value of $P(i, \gamma)$ is desirable as we want to minimize $|X_i(\gamma)|$. We consider the range of $r_{\text{upp}}$ as:

$$r_{\text{SO}}(0) \leq r_{\text{upp}} \leq r_{lF_0}. \tag{5.24}$$

Where $r_{lF_0}$ is the minimal congestion ratio found by PSO when applied to the whole set of failures, $F(= F_0)$. Otherwise, as $r_{SO}(0)$ is the minimal congestion ratio by definition of SO, $r_{\text{upp}}$ would be greater than $r_{lF_0}$ which is a trivial case as link duplication is not needed.

From Eq. (5.5), the range that should be considered for $\gamma$ is,

$$0 \leq \gamma \leq \gamma_{\text{max}}, \tag{5.25}$$

where

$$\gamma_{\text{max}} = \frac{r_{lF_0}}{r_{SO}(0)} - 1. \tag{5.26}$$
We evaluate the performance of the above three schemes by using four sample networks, as shown in Fig. 3.3. Traffic demands, link capacities are set as in LDO.

Both $r_l$ and $r_{SO}(0)$ are obtained with SO and PSO schemes. SO and PSO can be formulated as an Integer Linear Programming (ILP) problem [50]. However, as the network size grows, the ILP method cannot give results in a practical time. In [30], the authors presented a heuristic method to deal with larger networks. In the heuristic method, $I_{max}$ and $C_{max}$ which respectively represent the maximal number of initial weight sets and the maximal number of iteration per link weight are introduced. In our evaluation we use the heuristic method [30] with $I_{max}$ and $C_{max}$, respectively set to 1000 and 30 for faster convergence of the above three schemes within a practical time. $\gamma_{max}$ is determined by using Eq. (5.26).

For each scheme $i$, the protection ratios are evaluated for $\gamma$ varying from $\gamma_{max}$ to 0.00 with intervals of 0.10. Simulation results are shown from Table 5.2 to Table 5.5 for respectively networks 1, 2, 3, and 4. Due to time calculation constraints, only protection of at most three links are evaluated for the exhaustive search scheme under a given value of $\gamma$. Note that $|X_3(\gamma)|$ represents the number of protected links obtained by the exhaustive scheme when $\gamma$ is given.

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>LDO</th>
<th>E-LDO</th>
<th>Exhaus. Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.60</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.50</td>
<td>0.55</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>0.40</td>
<td>0.73</td>
<td>0.36</td>
<td>0.27</td>
</tr>
<tr>
<td>0.30</td>
<td>0.73</td>
<td>0.64</td>
<td>-</td>
</tr>
<tr>
<td>0.20</td>
<td>0.91</td>
<td>0.73</td>
<td>-</td>
</tr>
<tr>
<td>0.10</td>
<td>1.00</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-</td>
</tr>
</tbody>
</table>

From Table 5.2 to Table 5.5, we observe that the protection ratios increase when $\gamma$ decrease from $\gamma_{max}$ to 0.00. We also remark that, for a given particular value of $\gamma$, E-LDO shows a lower protection values compared to LDO in all
5. LINK-DUPLICATION MODEL FOR LINK WEIGHT OPTIMIZATION

Table 5.3: $\gamma$ values and protections ratios in network 2.

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>LDO</th>
<th>E-LDO</th>
<th>Exhaus. Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.54</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.44</td>
<td>0.23</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>0.34</td>
<td>0.41</td>
<td>0.27</td>
<td>-</td>
</tr>
<tr>
<td>0.24</td>
<td>0.77</td>
<td>0.68</td>
<td>-</td>
</tr>
<tr>
<td>0.14</td>
<td>0.91</td>
<td>0.86</td>
<td>-</td>
</tr>
<tr>
<td>0.04</td>
<td>1.00</td>
<td>0.95</td>
<td>-</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 5.4: $\gamma$ values and protections ratios in network 3.

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>LDO</th>
<th>E-LDO</th>
<th>Exhaus. Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.51</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.41</td>
<td>0.14</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>0.31</td>
<td>0.43</td>
<td>0.19</td>
<td>0.10</td>
</tr>
<tr>
<td>0.21</td>
<td>0.67</td>
<td>0.38</td>
<td>-</td>
</tr>
<tr>
<td>0.11</td>
<td>0.76</td>
<td>0.62</td>
<td>-</td>
</tr>
<tr>
<td>0.01</td>
<td>0.95</td>
<td>0.95</td>
<td>-</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-</td>
</tr>
</tbody>
</table>

examined network topologies. For network 3 and 4, protections ratios of E-LDO are very close if not, equal to those of the exhaustive scheme for the examined values of $\gamma$. In Eq. (5.5), $r_{upp} = (1 + \gamma) r_{SO}(0)$, means that $r_{upp}$ is $\gamma \times 100$ percent higher than $r_{SO}(0)$. In the following, we state our observations for some selected results to help readers to interpret our evaluation.

In network 1, for $\gamma = \gamma_{max} = 0.60$ the protection ratio is 0.00 for all schemes. This means that link protection is not needed if the network operator is able to manage a congestion ratio 60\% higher than the minimal congestion ratio. For $\gamma = 0.40$, we have 0.73, 0.36, 0.27 as protection ratios respectively for LDO, E-LDO and the exhaustive scheme. From Table 3.1, network 1 has 11 links, this means that to keep a manageable congestion ratio 40\% higher than the minimal congestion ratio, 8 ($=11 \times 0.73$) links need to be protected in LDO whereas only...
5.3 Enhanced link weight optimization scheme with link duplication

Table 5.5: \( \gamma \) values and protections ratios in network 4.

<table>
<thead>
<tr>
<th>( \gamma )</th>
<th>LDO</th>
<th>E-LDO</th>
<th>Exhaus. Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.11</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>1.01</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>0.91</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>0.81</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>0.71</td>
<td>0.10</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>0.61</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>0.51</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>0.41</td>
<td>0.20</td>
<td>0.20</td>
<td>-</td>
</tr>
<tr>
<td>0.11</td>
<td>0.95</td>
<td>0.75</td>
<td>-</td>
</tr>
<tr>
<td>0.01</td>
<td>1.00</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-</td>
</tr>
</tbody>
</table>

4 \((=11 \times 0.36)\) links should be protected in E-LDO and 3 \((=11 \times 0.27)\) links in the exhaustive scheme. In network 2, to keep a congestion ratio 34% higher than the minimal congestion ratio, 9 \((=22 \times 0.41)\) links have to be reinforced while 6 \((=22 \times 0.27)\) links would be enough in E-LDO. In network 3 as well, for a manageable congestion ratio 31% higher than the minimal one, protecting 9 \((=21 \times 0.43)\) links is required under LDO whereas only protecting 4 \((=21 \times 0.19)\) links is enough under E-LDO. Finally, in network 4, for \( \gamma = 0.50 \), the protection ratio of E-LDO matches that of the exhaustive scheme. Reinforcing 3 \((=20 \times 0.15)\) links is enough and sufficient as the exhaustive scheme examines all possible combinations of 3 links.

Overall, to keep the same manageable congestion ratio, E-LDO shows a lower protection ratio compared to LDO.

5.3.3 Computation time complexity

We evaluate and compare the computation time complexity of LDO and E-LDO.

In LDO, PSO is run only one time for every link protection where as in E-LDO, for one link protection it has to be run for the whole number of non-protected links. Let \( O(T(m)) \) be the computation time complexity of PSO when
it is applied to $H \subseteq F$ with $|H| = m$. Let $n$ represent the number of protected links.

For $n = 0$, in LDO, PSO is applied to the whole set $E$, $|E| = L$. The computation time complexity is therefore $O(T(L))$. However, in E-LDO PSO is applied $L$ times on, $E$ minus the protected single link. We have a computation complexity of $L \times O(T(L - 1))$.

For $n = 1$, in LDO, PSO is applied to $L - 1$ links as we consider only non-protected links failure. Therefore, $O(T(L - 1))$ is added to the computation time complexity found for $n = 0$. As PSO determines the optimal weight for each link, the computation time on $L$ links is higher than that on $L - 1$ links. The order of LDO’s computation time complexity is therefore $O(T(L))$. On the other hand, E-LDO will show a computation time complexity being the sum of $(L - 1) \times O(T(L - 2))$ and $L \times O(T(L - 1))$. For the same reason mentioned above, the complexity is $L \times O(T(L - 1))$.

The worst possible case is the situation where the manageable congestion ratio is met only after all the links are protected. That is when $n = L$. This may happen when $r_{up} = r_{SO}(0)$. We consider this case for both schemes. Under this case, higher values of $n$ shall be considered. For higher values of $n$, the same observation as above is applied as the dominant terms are unchanged. We can say that for $n = L$ LDO and the E-LDO have respectively $O(T(L))$ and $L \times O(T(L - 1))$ as computation time complexity.

The above observation can be generalized as it is verified for all $n$ from 1 to $L$. Furthermore, $O(T(L - 1))$ is bounded by $O(T(L))$ as computation time on $L$ links is higher than that on $L - 1$ links. In term of computation time complexity upper bound, E-LDO computation time complexity is at most $L$ times larger than that of LDO. The summary of this comparison is shown in Table 5.6.

5.4 Summary

We proposed a link-duplication model, which we call an LD model, for link weight optimization to reduce link duplication or protection in IP networks by choosing an appropriate set of link weights with the condition of maintaining the manageable congestion under any single-link failure scenario. For that purpose two
5.4 Summary

Table 5.6: Comparison of computational complexities.

<table>
<thead>
<tr>
<th>Number of protected links ($n$)</th>
<th>LDO</th>
<th>E-LDO</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$O(T(L))$</td>
<td>$L \times O(T(L - 1))$</td>
</tr>
<tr>
<td>1</td>
<td>$O(T(L))$</td>
<td>$L \times O(T(L - 1))$</td>
</tr>
<tr>
<td>2</td>
<td>$O(T(L))$</td>
<td>$L \times O(T(L - 1))$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$</td>
<td>E</td>
<td>=L$</td>
</tr>
</tbody>
</table>

schemes were introduced. The first scheme, which we call link-duplication optimization (LDO), uses PSO to find the suitable link weight set but optimizes link weights before determining the link to protect. This is an issue as some protection patterns with a lower number of links, that yield the same performance might be skipped. The second scheme, which we call Enhanced LDO (E-LDO), corrects that weakness by optimizing link weights after considering multiple link protection candidates. For each candidate, link weights are optimized so as to reduce the worst failure congestion ratio. The protection candidate that shows the lowest congestion ratio is chosen as the desired protection pattern. We compared the performance of both LDO and E-LDO in terms of the required number of link protections with the condition of keeping a manageable congestion ratio under both failure and non failure scenarios. E-LDO reduces the number of link protections compared to LDO.
5. LINK-DUPLICATION MODEL FOR LINK WEIGHT OPTIMIZATION
Chapter 6

Comparison of both penalty-aware and link-duplication models when considering additional resources

This chapter compares the penalty-aware (LA) model and the link-duplication (LD) aware model. Since the LD model considers additional resources (duplication of links), a fair comparison with the PA model would require a consideration of additional capacity as well. In this section, we consider S-PSO as the scheme that represents the PA model as it has the smallest penalty under the minimized worst congestion ratio. In Section 6.1 for the PA model, a mathematical formulation of additional capacity required to meet the manageable congestion ratio is presented. Evaluation and comparison of the required additional capacity for both models is shown in Section 6.2. Finally in Section 6.3 we summarize the results and observations.
6. COMPARISON OF BOTH PENALTY-AWARE AND LINK-DUPLICATION MODELS WHEN CONSIDERING ADDITIONAL RESOURCES

6.1 Mathematical formulation of penalty-aware model considering additional resources

In the PA model, capacity is added to the network after failure occurs in order to reduce the increased congestion ratio. Therefore, designing an algorithm that reduces backup, or additional capacity, while keeping a manageable congestion ratio should be investigated for cost reduction.

Let us suppose that the manageable congestion ratio at network operation start-time is $r_{\text{upp}}$ and the congestion ratio under worst case failure is $r_{l_F0}$. With the PA model we want to answer the following question: how much additional capacity is required to keep $r_{\text{upp}}$ even under the worst case failure without changing routing parameters such as link weights?

As defined in Section 2.1, $c_l$ represents the capacity of $l \in E$. Let us define $k$ by,

$$k = \frac{r_{l_F0}}{r_{\text{upp}}}. \quad (6.1)$$

$k$ represents the capacity utilization growth factor between the state where the congestion ratio is manageable and the state where we have the worst congestion ratio. Therefore, increasing the capacity $c_l$ of each link $l$ to $k \times c_l$ will keep the congestion to a value equal to $r_{\text{upp}}$ even under the worst case failure. As a result the additional capacity per link is $c_{lad} = (k - 1) \times c_l$ and the total additional capacity is expressed by,

$$C_{ad} = \sum_{l \in E} c_{lad} = (k - 1) \sum_{l \in E} c_l. \quad (6.2)$$

Let us define $C_0 = \sum_{l \in E} c_l$. We have,

$$C_{ad} = \left(\frac{r_{l_F0}}{r_{\text{upp}}} - 1\right) \times C_0. \quad (6.3)$$

From Eq. (5.26) we have,

$$C_{ad} = \left(\frac{1}{1 + \gamma r_{SO(0)}} - 1\right) \times C_0. \quad (6.4)$$
6.2 Evaluation and comparison of both penalty-aware model and the link-duplication model

We compare the PA model and the LD model. The comparison metric employs the additional resource ratio $A_r$, which is defined by,

$$A_r = \frac{C_{ad}}{C_0}. \quad (6.5)$$

As $C_0$ is a given value, $A_r$ is obtained by evaluating $C_{ad}$. For the PA model, $C_{ad}$ is evaluated by Eq. (6.4). For the LD model it is evaluated by,

$$C_{ad} = \sum_{l \in X_2} c_l, \quad (6.6)$$

where $X_2$ is the set of protected links found in E-LDO.

Evaluation conditions including traffic demands, link initial capacities and sample networks are the same as those chosen in Section 5.2.2. $r_{F0}$ and $r_{SO(0)}$ are calculated and $\gamma$ is selected as mentioned in Section 5.2.2 as well. For different values of $\gamma$, the evaluation results are shown from Table 6.1 to Table 6.4 for networks 1, 2, 3, and 4, respectively.

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>LD model</th>
<th>PA model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.60</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.50</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>0.40</td>
<td>0.36</td>
<td>0.14</td>
</tr>
<tr>
<td>0.30</td>
<td>0.64</td>
<td>0.23</td>
</tr>
<tr>
<td>0.20</td>
<td>0.73</td>
<td>0.33</td>
</tr>
<tr>
<td>0.10</td>
<td>1.00</td>
<td>0.45</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Evaluation results show that for values of $\gamma$ closed to $\gamma_{max}$ the LD model has an additional resource ratio lower or equal to that of the PA model. This can be verified for $\gamma = 0.50, 0.44, 0.41$ and 1.01, respectively in networks 1, 2, 3 and 4. On the other hand, for smaller values of $\gamma$, the PA model has a lower additional
6. COMPARISON OF BOTH PENALTY-AWARE AND LINK-DUPLICATION MODELS WHEN CONSIDERING ADDITIONAL RESOURCES

Table 6.2: Additional resource ratios for $\gamma$ in network 2.

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>LD model</th>
<th>PA model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.54</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.44</td>
<td>0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>0.34</td>
<td>0.28</td>
<td>0.15</td>
</tr>
<tr>
<td>0.24</td>
<td>0.68</td>
<td>0.24</td>
</tr>
<tr>
<td>0.14</td>
<td>0.87</td>
<td>0.35</td>
</tr>
<tr>
<td>0.04</td>
<td>1.00</td>
<td>0.48</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Table 6.3: Additional resource ratios for $\gamma$ in network 3.

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>LD model</th>
<th>PA model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.51</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.41</td>
<td>0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>0.31</td>
<td>0.20</td>
<td>0.15</td>
</tr>
<tr>
<td>0.21</td>
<td>0.39</td>
<td>0.24</td>
</tr>
<tr>
<td>0.11</td>
<td>0.61</td>
<td>0.36</td>
</tr>
<tr>
<td>0.01</td>
<td>0.95</td>
<td>0.49</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>0.51</td>
</tr>
</tbody>
</table>

resource ratio than that of the LD model in networks 1, 2 and 3. The PA model adds the exact amount of capacity required in the whole network whereas the LD model duplicates links thus doubles the capacity at the protected place. This leads to an increase of unnecessary capacity. As $\gamma$ decreases, the number of protected links grows. This increases even more the unnecessary capacity and explains the performance difference between the LD model and the PA model.

However, in network 4, the LD model outperforms the PA model even for small values of $\gamma$. This is because the gap between the worst congestion ratio and the minimal congestion ratio is large. The worst congestion ratio is more than the double of the minimal congestion ratio (in Eq. (5.26) replace $\gamma_{\text{max}}$ with 1.11). For small values of $\gamma$, that is, when the manageable congestion $r_{\text{upp}}$ is close to the minimal congestion ratio, more than two times the initial capacity is required.
6.2 Evaluation and comparison of both penalty-aware model and the link-duplication model

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>LD model</th>
<th>PA model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.11</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>1.01</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>0.91</td>
<td>0.05</td>
<td>0.10</td>
</tr>
<tr>
<td>0.81</td>
<td>0.10</td>
<td>0.17</td>
</tr>
<tr>
<td>0.71</td>
<td>0.10</td>
<td>0.23</td>
</tr>
<tr>
<td>0.61</td>
<td>0.15</td>
<td>0.31</td>
</tr>
<tr>
<td>0.51</td>
<td>0.15</td>
<td>0.39</td>
</tr>
<tr>
<td>0.41</td>
<td>0.21</td>
<td>0.49</td>
</tr>
<tr>
<td>0.11</td>
<td>0.76</td>
<td>0.90</td>
</tr>
<tr>
<td>0.01</td>
<td>1.00</td>
<td>1.09</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>1.11</td>
</tr>
</tbody>
</table>

to keep $r_{\text{upp}}$ under the worst case failure. Therefore, the PA model requires an additional capacity that is larger than the initial capacity.

Now, for the LD model we assume that a protected link does not fail. Therefore, protecting all the links always guarantee the minimal congestion ratio. As a result, in any network including network 4, the LD model requires at most, an additional capacity equal the initial capacity. This shows why the LD model has a lower additional capacity ratio than that of the PA model.

The gap between the worst congestion ratio and the minimal congestion in network 4, can be explained by the fact that it has a higher average node degree compared to other sample networks (see Table 3.1). The degree of a node is the number of nodes adjacent to that node. The average node degree in a network is evaluated by $\frac{|E|}{|V|}$. In [30], it was shown that the gap between the worst congestion ratio and the minimal congestion ratio increases with the average network degree. From this fact we can say that the LD model is suitable in terms of additional capacity ratios for networks with a high average node degree.

From the network operator point of view, traffic routes shall not be changed in some cases due to latency requirements or geographically restriction of specific network services or functions. In the LD model traffic routes are not affected as only switching from the primary link to the backup link is enough to tackle link
failures. In the PA model, this is not the case since capacity is just added to accommodate the rerouted traffic after failure.

In summary, network operators should consider the network characteristics, the requirements of latency and continuity for traffic and the geographical restriction of services to choose a suitable protection model against failure.

6.3 Summary

We evaluated and compared both the PA and LD models. Since additional in resources is considered for the LD model we incorporate additional capacity in the PA model. For the PA scheme based on an approach that adds capacity to the network when failure occurs in order to keep the manageable congestion ratio. For the PA model we introduce a mathematical formulation that aims to determine the minimal additional capacity to provide in order to maintain the manageable congestion under any single-link failure scenario. We then compare the LD model to the PA model that incorporates additional capacity features. Evaluation results show that the performance difference between the LD model and this PA model in terms of the required additional capacity depends on the network characteristics. The requirements of latency and continuity for traffic and geographical restriction of services should be taken into consideration when deciding which model to use.
Chapter 7

Conclusions and future works

7.1 Conclusions

Our objective is to determine a link weight set that keeps network congestion ratio below the manageable threshold for every single-link failure. After failure occurs, the congestion ratio rises. This restricts addable traffic and increases packet drop rate. A possible way to tackle this problem could be increasing the capacity of the network. However, due expenditure constraints a reduction of the additional capacity is required.

Meanwhile, communication paths are determined by link weights that are set by network operators. Since these links can be tuned to diverted traffic, they can also be optimized to reduce the congestion ratio thus reduce the required additional capacity to match the manageable congestion ratio.

In the basic model of link weight optimization [30], the focus is to minimize the additional capacity. In this model, preventive start-time optimization (PSO) is presented. PSO determines a set of link weights that minimizes the worst case congestion ratio under any single-link failure scenario. However, applying PSO link weight sets in the network when there is no failure may cause a larger congestion ratio compared to the conventional scheme. This issue was mentioned in [30] but was not addressed. We have addressed this problem since in networks with few failures, that burden would be carried all along.

We have proposed a penalty-aware (PA) model for link weight optimization to eliminate that penalty while reducing the worst case congestion ratio. In our
PA model we presented two schemes: preventive start-time optimization with no penalty (PSO-NP) and strengthened preventive start-time optimization (S-PSO). PSO-NP scheme is based on the SO (start-time optimization) scheme. SO determines a link weight set that minimizes the congestion ratio under a non-failure scenario. This set may not be unique and there are actually multiple sets obtained by SO that would produce the same result. PSO-NP relies on choosing among the SO-generated sets, the set best prepared to deal with the worst-case congestion in our network. Contrary to PSO-NP, S-PSO is based on PSO but relies on choosing among the PSO-generated sets the set that shows the lowest penalty. In that sense PSO-NP eliminates the penalty while reducing the worst congestion whereas S-PSO minimizes the worst congestion ratio as PSO does while reducing the penalty compared to PSO.

A direct ILP-based mathematical formulation as well as heuristic formulation of both PSO-NP and S-PSO were described. Due to network size constraints only the Heuristic formulation was used to evaluate the performance of both scheme. Numerical results have shown that PSO-NP reduces PSO penalty in case of no failure, while significantly reducing the congestion ratio under the worst-case failure. Also S-PSO was shown to reduce the penalty while minimizing the worst-case congestion ratio. We have also considered the trade-off between the penalty when there is no failure and the reduction ratio of the worst case congestion. We have explored that in a general preventive start-time optimization (GPSO). GPSO determines the weight set that minimizes the worst-case congestion ratio when the penalty under no failure is given. We have shown that GPSO includes previous start-time optimization schemes when the penalty bound is appropriately set. GPSO provides a flexible weight set that considers both failure and non-failure scenarios.

GPSO is effective in finding a link weight set that reduces the congestion under both failure and non-failure cases. However, it does not guarantee the manageable congestion as it considers penalty. To tackle this problem, in this thesis we propose a link-duplication (LD) model to suppress link failure in the first place in order to always meet the manageable congestion. For this purpose we consider the duplication or reinforcement of links which is broadly used to make network reliable. Link duplication provides fast recovery as only switching from
the failed link to the backup link will hide the failure at upper layers. However, due to capital expenditure constraints, not every link can be duplicated. Giving priority to some selected links makes sense. As mentioned above, communication paths are determined by link weights that are configured in advance. Therefore, choosing an appropriate set of link weights may reduce the number of links that actually need to be duplicated in order to keep a manageable congestion under failure and also non failure cases. Now, PSO also determines the link failure which creates the worst congestion after failure. Since by duplicating this link we can assume it no more fails, the LD model uses PSO to find the smallest number of links to protect so as to guarantee a manageable congestion under any single-link failure. In the LD model we have explored two schemes. The first scheme which we call link-duplication optimization (LDO) optimizes link weights before determining the link to protect. This is an issue as some protection patterns that yield the same performance with a lower number of links might be skipped. We solved that issue by enhancing LDO in a second scheme we denoted E-LDO. E-LDO optimizes link weights after considering multiple link protection candidates. For each candidate, link weights are optimized so as to reduce the worst failure congestion ratio. The protection candidate that shows the lowest congestion ratio is chosen as the desired protection pattern. We compared the performance of these LD model based schemes in terms of the required number of link protections with the condition of keeping a manageable congestion ratio under any single-link failure scenario.

In the third part of this thesis we compare the PA model and the LD model. Since the LD model considers additional resources, a fair comparison with the PA model would require considering additional capacity in the PA model as well. We incorporate additional capacity in the PA model. For the PA model we introduce a mathematical formulation that aims to determine the minimal additional capacity in order to maintain the manageable congestion under any single-link failure scenario. We then compare the LD model to the PA model that incorporates additional capacity features. Evaluation results show that the performance difference between the LD model and this PA model in terms of the required additional capacity depends on the network characteristics. The requirements of
7. CONCLUSIONS AND FUTURE WORKS

latency and continuity for traffic and geographical restriction of services should be taken into consideration when deciding which model to use.

7.2 Future works

The work presented in this thesis opens the ways to several directions for future work. As future work, the formulation and performance evaluation of both PA and LD models under traffic uncertainty as well as including layers such as the optical one for the LD model shall be considered.

In this thesis we considered a situation where traffic demand is known but in actual commercial network traffic is uncertain [8, 10, 12, 37, 52, 53]. In fact traffic highly varies from day to night in the suburbs of big cities as people move from home to their work place. It also changes due to more unpredictable factors such as disasters [13, 43, 47], software updates, games or movie releases, invents and so on. These factors tremendously increase traffic demand at a specific place and time [26, 45, 69]. In the basic model of link weight optimization, traffic uncertainty was addressed [63, 64]. As a next step traffic uncertainty in the PA model and its related schemes shall be considered. In the LD model too, traffic unpredictability should be considered when pinpointing the links that need to be duplicated.

Moreover, the LD model and its dependent schemes can be expanded to include the optical layer as we only focused on the data link layer and the IP layer for link duplication. By considering the optical layer, additional resources can be inspected at a lower granularity since bundle of optical links constitute a link a the data link layer. This may lead to a reduction of the required additional resources as added discrete resources become less sparc e and resource sharing among links of the same bundle can envisaged.

However at the optical layer, shared risk link groups (SRGs) [44, 57] shall be taken into consideration [49, 56, 58]. SRG is a group of link that share a common resource or function such that when malfunction occurs all the related links fail. In the optical layer bundled links are grouped together in a span [1]. This increases the risk of having all links failed if a span cut occurs. This kind of event may happen during disasters and large scale network attacks and affects
network operation \cite{14, 16, 24, 39}. PA and LD models should be upgraded to also deal with simultaneous multi-link failures as this thesis addresses only single link failures. The work in \cite{28} considers an optimization problem to reduce the backup resources under random and simultaneous link failures by introducing the problematic survivability. PA and LD models should be extended to deal with random and simultaneous link failures to keep the problematic survivability.

Furthermore, disasters and network attacks are less likely to happen compare to traffic spikes or software updates \cite{7, 21, 59}. In that sense simultaneous link failures rarely happen. A permanent protection against these events will result in costly networks in an era where data related revenue is decreasing with the spread of Mobile Virtual Network Operators (MVNOs) \cite{5, 11, 20, 35, 40}. There is a need to include probabilistic factors \cite{57} when designing new PA and LD models in order to deal with random and simultaneous link failures at a lower cost.

Finally, PA and LD models should be investigated and compared in the framework of network function virtualization (NFV) and software defined networking (SDN). NFV decouples network functions from proprietary hardware to enable network operators to easily scale resources on demand at any given location above a cheaper hardware. SDN makes it possible to dynamically set or update communication paths based on specific predefined events \cite{66}. In this thesis we did not consider the factors that NFV and SDN bring to the table in terms of network traffic engineering. Since NFV/SDN is gradually being introduced in commercial networks \cite{6}, it should be taken into consideration when investigating the applicability of PA and LD models.
7. CONCLUSIONS AND FUTURE WORKS
Bibliography


[40] C.M. Machuca, T. Krauss, A. Basta, E. Grigoreva, and W. Kellerer. Alternatives for the automotive industry to become an mvno. In *International Conference on Transparent Optical Networks (ICTON)*, 2014. 69


[59] L. Pang, J. Zhou, and D. Liu. The probability prediction of storm surge disaster along south-east coasts of china. IEEE, 2011. 69

76


Publications

List of Publications related to the thesis

Journal Papers


International Conference Papers

National Conference Papers

