Engineering Optimization
Vol. 00, No. 00, September 2016, 1–23

RESEARCH ARTICLE

A network-wide exact optimization approach for multiobjective routing with path protection in multiservice Multiprotocol Label Switching networks

Rita Girão-Silva[a,b*], José Craveirinha[b], Teresa Gomes[a,b], Lúcia Martins[a,b], João Clímaco[b], João Campos[a,b,c]

[a]*DEEC-FCTUC; Pólo II, Pinhal de Marrocos; P-3030-290 Coimbra; Portugal*; [b]*INESC-Coimbra – DEEC-FCTUC; Pólo II, Pinhal de Marrocos; P-3030-290 Coimbra; Portugal*; [c]*Currently at CERN, Switzerland*

(*September 5, 2016*)

A multiobjective routing model for Multiprotocol Label Switching networks with multiple service types and path protection is presented in this paper. The routing problem is formulated as a biobjective integer program, where the considered objectives are formulated according to a network-wide optimization approach, i.e. the objective functions of the route optimization problem depend explicitly on all traffic flows in the network. A disjoint path pair is considered for each traffic trunk, which guarantees protection to the associated connection. A link-path formulation is proposed for the problem, in which a set of possible pairs of paths is previously devised for each traffic trunk. An exact method (based on the classical constraint method for solving multiobjective problems) is developed for solving the formulated problem. An extensive experimental study, with results on network performance measures in various randomly generated networks, is also presented and discussed.

Keywords: routing models; multiobjective optimization; Multiprotocol Label Switching networks; path protection

*Corresponding author. Email: rita@deec.uc.pt

ISSN: 0305-215X print/ISSN 1029-0273 online

DOI: 10.1080/0305215YYxxxxxxxx
http://www.informaworld.com

1. Introduction

In modern multiservice networks, different classes of traffic with different service requirements are considered. The routing calculation and optimization problems in these networks can be quite challenging, because of the multi-dimensional and often conflicting nature of the performance and cost metrics. In communication networks, routing models are focused on the selection of a sequence of network resources (constituting paths or 'routes') aiming at the optimization of some route related objective functions, while satisfying a set of multiple and heterogeneous Quality of Service (QoS) constraints. The performance of different routing decisions may be evaluated, taking into account the chosen route related network metrics.

The formulation of routing problems in these types of networks as multiple objective optimization problems is potentially advantageous. This allows for the evaluation, in a mathematically consistent form, of the trade-offs among distinct and potentially conflicting performance metrics and/or other network cost function(s).

In multiobjective optimization problems, optimal (ideal) solutions (i.e. solutions which optimize all the objective functions simultaneously) are usually unfeasible. Therefore, in these problems, the general purpose is to find non-dominated (or Pareto optimal) solutions (Steuer 1986) and select one of those 'trade-off solutions'. The set of non-dominated solutions of a multiobjective optimization problem constitutes the so-called Pareto front. A non-dominated solution is a feasible solution such that it is not possible to find another feasible solution with better value of an objective function without worsening the value of at least one of the other objective functions. A weakly non-dominated solution is a feasible solution such that it is not possible to find another solution which improves the values for all the objective functions. Weakly non-dominated solutions are not non-dominated solutions. Note that non-dominated solutions may be supported or unsupported. Supported solutions are non-dominated solutions for which the corresponding objective function values are located on the boundary of the convex hull of the Pareto front.

Regarding state of the art reviews on multicriteria routing optimization models for modern communication networks, see (Clímaco *et al.* 2016, 2007). A proposal of a conceptual framework for consistent multicriteria routing approaches in QoS-IP (Internet Protocol) networks, is given in (Wierzbicki and Burakowski 2011).

1.1. *Path protection in MPLS networks*

In Multiprotocol Label Switching (MPLS) networks, packets are forwarded through Label Switched Paths (LSPs), according to certain technical rules. An important problem in traffic engineering is to provide protection to a connection, which can be achieved with the establishment of a disjoint path pair for each traffic trunk (aggregation of traffic flows between a pair of nodes/routers of the same Forward Equivalence Class – FEC) on the network. This procedure is known as path protection and it entails the establishment of a node (preferably) or a link-disjoint pair of LSPs.

Path protection can be used in order to guarantee the existence of a path to carry the offered traffic, in situations of single fault. The definition of a full path protection mechanism includes the protection configuration, the fault detection, the fault notification, and the protection switching (Huang *et al.* 2002). This work focuses solely on the protection configuration and assumes the other aspects are covered by appropriate MPLS protocols (Sharma and Hellstrand 2003).

A scheme of dedicated path protection is considered. In the normal conditions of a network, i.e. in the absence of faults, the active path (also designated as working path or primary path) is used to carry the offered traffic; whenever a fault occurs, the backup path (or secondary path) is used to carry that traffic. To guarantee the reliability of the end-to-end connections, this pair of paths must not have a common risk of failure. Note that both paths must have a provisioned capacity so that they can carry the designed traffic, which requires the reservation of an adequate amount of bandwidth in both paths. Therefore, this is a 1+1 dedicated path protection scheme (Huang *et al.* 2002).

Protection can be local or global, and resource or path oriented. With local protection, the purpose is to protect against a link or node fault and to minimize the time required for fault recovery. The goal of global protection is to protect against any link or node fault in a path. A protection mechanism is resource oriented if it tries to protect a particular network element (link and/or node), or path oriented if the protection mechanism is focused on the protection of a particular active path (Jorge and Gomes 2006).

The most widely known local protection mechanism in MPLS is Fast Reroute (FRR). In the event of a failure, this local protection scheme guarantees that traffic may be redirected to a preset backup LSP which originates at a Point of Local Repair (PLR) and merges with the active LSP at a specified Merge Point. Two different local protection techniques are specified by MPLS FRR: facility backup (a resource-oriented protection mechanism) and one-to-one backup (a path-oriented protection mechanism) (Alvarez 2006, Jorge and Gomes 2006, Pan *et al.* 2005). The facility backup method creates a bypass tunnel to protect a potential failure point and it uses label stacking to reroute multiple protected LSPs using a single backup LSP. The one-to-one backup method does not use label stacking and it creates detour LSPs for each protected LSP at each potential PLR. Although other local protection schemes were earlier proposed by different authors, FRR in MPLS networks has become a standard.

Many papers have been published concerning global protection (i.e. end-to-end path protection) specifically in the context of MPLS networks, and also for other types of networks, that could easily be adapted to MPLS networks. In (Huang *et al.* 2002), a simple, scalable, fast, and efficient path protection mechanism is proposed. The authors focus on the control of the delay that must be incurred by the notification message traveling from the fault detection node to the protection switching node, and attempt to minimize that delay by using a special tree structure to efficiently distribute fault and/or recovery information. In (Naraghi-Pour and Desai 2008), the authors propose a new approach for Virtual Private Network (VPN) traffic engineering with path protection in MPLS networks carrying QoS and Best Effort traffic. This approach deals with the existence of path cycles and guarantees the control of the maximum path length and of the size of the label space in each Label Switched Router. In (Wang and Ergun 2005), the authors propose a technique which might not be considered a protection scheme in a strict sense, as the backup resources are not reserved prior to the failure. The backup resources are merely identified prior to the failure and once a failure occurs, the technique consists of finding a usable backup path. The authors claim that their model is significantly more cost-efficient than classical backup path reservation. There are other approaches based on multiple backup paths. In (Alouneh *et al.* 2009), the same information is sent across multiple LSPs. At the destination node, the information regarding the IP packet is retrieved from the information that actually arrived at that node coming from the different backup paths. This approach is further extended in (Mobasheri and Moghadam 2014), where the incoming demands are assumed to come from sub-networks of a smart grid entering the IP/MPLS backbone.

1.2. *Description of this work*

The aim of this work is the development and exact resolution of a multiobjective routing model with path protection, which involves the simultaneous calculation of pairs of node (preferably) or link-disjoint LSPs for all node-to-node traffic flows, where the routing optimization problem is formulated in terms of a network-wide optimization approach, in the sense defined in (Mitra and Ramakrishnan 2001) and in (Craveirinha *et al.* 2008). Therefore, the objective functions of the route optimization problem must depend explicitly on all traffic flows in the network, allowing for an exact representation of the interactions and interdependencies among all traffic flows (unlike in flow-oriented optimization routing models).

An earlier work by the authors (Girão-Silva *et al.* 2015) addressed a network-wide optimization model for multiobjective routing in multiservice MPLS networks. Notice however that in the model in (Girão-Silva *et al.* 2015) there is the possibility of traffic splitting (i.e. each traffic flow may be divided by different paths from one source node to a destination node), but not path protection. The optimization problem structure is similar but the actual problem addressed in this paper is clearly different, in terms of one of the objective functions and the problem constraints.

The routing problem considered here is formulated as a biobjective network flow programming optimization model. The considered objectives are the minimization of the bandwidth transport cost and the minimization of the load cost in the network links, which are similar to the objective functions in (Erbas and Erbas 2003) and (Girão-Silva *et al.* 2015). A link-path formulation of the optimization problem is considered, in which a set of possible pairs of paths is previously devised for each traffic trunk. The actual process of finding adequate non-dominated path pairs (among those in the provided set) for all the traffic flows is performed by an exact method. In multiple criteria/multiobjective optimization, an exact method is one that enables the exact calculation of all non-dominated solutions of the problem. In (Cohon 1978), several classical exact methods are presented for solving multiobjective optimization problems. In (Ehrgott and Gandibleux 2000), an annotated bibliography of multiple objective combinatorial optimization problems is provided, including a section on available exact methods.

The exact resolution method applied to this routing problem is based on the classical constraint method (Cohon 1978) and was already considered in (Girão-Silva *et al.* 2015). Some important features of the adaptation performed on the classical constraint method are: i) specific strategies to avoid weakly non-dominated solutions; ii) choice of a specific area of the objective function space to be explored and where interesting solutions may be found; iii) choice of a specific solution as an adequate compromise solution to the routing problem. The latter features rely on a system of preferences established through the definition of preference regions in the objective function space. These regions are obtained from aspiration and reservation levels (preference thresholds) defined for the considered objective functions.

An experimental study is presented in this paper, regarding experiments carried out with a set of randomly generated networks with topologies obtained with the gt-itm (Georgia Tech – Internetwork Topology Models) software (GT-ITM 2000). Relevant network performance measures, based on the ones presented in (Srivastava *et al.* 2005), were calculated and a statistical study on their average values and variation ranges was performed. Some of the performance measures are related to the links utilization, which is a very important parameter concerning network traffic carrying capabilities. Different global routing solutions can be compared in terms of the objective function values and also of these network performance measures. A comparison of the final solutions with the ideal solutions that would be obtained if single objective problems were solved for each objective function, is also performed.

The main contributions of this paper are the following:

- formulation of a biobjective routing model for MPLS networks with path protection and multiple service types. The considered objectives (the minimization of the bandwidth routing cost and the minimization of the load cost in the network links) and constraints are formulated according to a network-wide optimization approach;
- use of an exact method to solve the aforementioned model (rather than heuristic or meta-heuristic techniques) in the context of a link-path formulation of the routing problem. The resolution model has two stages: i) a set of feasible path pairs is devised; ii) a specific path pair is chosen for each traffic flow, according to a model based on the classical constraint method;
- an experimental study, where a double randomness is considered: not only are the MPLS networks topologies randomly generated, but also the matrix of the traffic offered to those networks is randomly generated;
- analysis of the results obtained for relevant network performance measures related to the links utilization and also for other performance measures that allow for a comparison of the obtained solutions with solutions from single objective optimization models.

The paper is organized as follows: in the next section, the routing model and the mathematical formulation of the biobjective optimization problem are described. In section 3, the proposed method for solving the formulated problem, is revised. A fourth section describing the considered experimental study and the obtained results is followed by a final section with some conclusions and an outline of future work. Some supplementary material to this paper is available at [LINK MISSING], including some further explanation regarding the calculation of candidate path pairs, the pseudo-code description of the used algorithm and some further experimental results.

2. Model description

The notation of the model and the biobjective mathematical programming formulations of the biobjective routing model for MPLS networks with path protection, are described next.

2.1. *Nomenclature and notation*

A network $(\mathcal{N}, \mathcal{A})$ with directed arcs (or links) is considered, where $\mathcal{N} = \{1, \cdots, |\mathcal{N}|\}$ is the set of nodes in the network and $\mathcal{A} = \{1, \cdots, |\mathcal{A}|\}$ is the set of links. Throughout the paper, the network nodes will be identified by indices $i, j, v \in \mathcal{N}$ and the network links by index $k \in \mathcal{A}$. If the directed link k represents the link originating in i and destined for j, then k' will represent the symmetrical link, i.e. the link originating in j and destined for i. Notice that if a link k exists, then its symmetrical k' will exist as well. The capacity of each network link is given by u_k [Gbit/s], $k \in \mathcal{A}$, with $u_k \neq 0$. The problem is formulated under the more general assumption that there may be $u_k \neq u_{k'}$, i.e. a pair of links k and k' may have assymmetrical capacities.

Path protection must be provided in the context of the proposed routing model. For this reason, the path pairs for each origin-destination pair must be node (preferably) or link-disjoint. If no disjoint path pairs exist for a specific origin-destination pair, then this particular traffic flow will not be considered in the optimization problem to be solved and will be handled separately, since path protection cannot be provided in this case. This is beyond the scope of this work.

Different types of services are considered, with the set of services in the network being denoted by $\mathcal{S}$. Different path pairs may be obtained for different flows originating in i and destined for j, according to the service they belong to. The bandwidth required by flow $t \equiv (i,j,s)$ (i.e. the traffic flow associated with service $s \in \mathcal{S}$ and originating in i and destined for j) is $d_t = d_{(i,j,s)}$. The network demands may be assymmetrical, i.e. there may be $d_t \neq d_{t'}$, where $t \equiv (i,j,s)$ and $t' \equiv (j,i,s)$.

The set of offered flows is $\mathcal{T}$. The subset $\mathcal{T}_N$ includes the flows $t \in \mathcal{T}$ for which there is at least a node-disjoint path pair (with certain constraints – see sub-section 2.2.1); the subset $\mathcal{T}_A$ includes the flows $t \notin \mathcal{T}_N$ for which there is at least a link-disjoint path pair (with certain constraints – see sub-section 2.2.2).

Let $p_{t,l}^m$ be the l-th feasible path pair for flow $t \in \{\mathcal{T}_N \cup \mathcal{T}_A\}$, where $m = 1$ refers to the active path and $m = 2$ refers to the backup path, and $l = 1, \cdots, L_t$. The set of disjoint path pairs that is considered for each flow t is $\mathcal{L}_t$ and its size is $L_t = |\mathcal{L}_t|$. A path may be identified by its nodes, i.e. $p_{t,l}^m =< i, i_1, i_2, \cdots, j >$, or by its links, i.e. $p_{t,l}^m =< k_1, k_2, \cdots, k_M >$. Notice that no parallel links are possible, so a path may be fully identified by its sequence of links or by its sequence of nodes. The paths for flows $t \equiv (i,j,s)$ and $t' \equiv (j,i,s)$ must be topologically symmetrical, i.e. for each path originating in i and destined for j there is a symmetrical path (originating in j and destined for i). If $p_{t,l}^m =< i, i_1, i_2, \cdots, j >$, then $p_{t',l}^m =< j, \cdots, i_2, i_1, i >$ (or, in terms of links, if $p_{t,l}^m =< k_1, k_2, \cdots, k_M >$, then $p_{t',l}^m =< k'_M, \cdots, k'_2, k'_1 >$).

There is the possibility of bandwidth overbooking (since this is a common practice by network operators), expressed by a factor which varies according to the service and to the link to be considered. If a flow of service s offered to link k has an overbooking factor ξ_{ks}, this is equivalent to considering that the capacity occupied in the link by that flow is the product of the nominal bandwidth of the flow (d_t) by ξ_{ks}.

2.2. *Calculation of candidate path pairs*

In the first stage of the resolution approach, a set of candidate path pairs is obtained for each traffic flow t, in the context of the proposed link-path formulation.

This stage begins with the enumeration of all the flows $t \in \mathcal{T}_N$ and $t \in \mathcal{T}_A$. The remaining flows $t \in \mathcal{T} \setminus \{\mathcal{T}_N \cup \mathcal{T}_A\}$ will not be part of the problem to be solved and will be handled separately. Afterwards, an enumeration algorithm will be used so that feasible disjoint path pairs are obtained for each flow, as described next.

In the next two subsections, the following notation will be used: a feasible path pair for flow t is designated by p_t^m, where $m = 1$ refers to the active path and $m = 2$ refers to the backup path; a binary value $a_t^{k,m}$ is defined and it is 1 if link $k \in p_t^m$ and 0 otherwise, with $m = 1$ or 2, $k \in \mathcal{A}$ and $t \in \{\mathcal{T}_N \cup \mathcal{T}_A\}$; $\delta(v^+)$ represents the set of links emerging from node $v \in \mathcal{N}$ and $\delta(v^-)$ represents the set of links incident on node $v \in \mathcal{N}$.

2.2.1. *Enumeration of the flows with node-disjoint path pairs*

The following optimization problem is solved for each flow $t \equiv (i,j,s) \in \mathcal{T}$, with $i < j$, for enumerating flows with node-disjoint path pairs and a maximal number of links per path. Notice that it is sufficient to solve the problem for flows $t \equiv (i,j,s) \in \mathcal{T}$, with $i < j$, as it is assumed that a specific link always has a symmetrical link. Therefore, if flow $t \equiv (i, j, s) \in \mathcal{T}$, with $i < j$, has a node-disjoint path pair, there will surely be a topologically symmetrical path pair, which is also node-disjoint, for flow t'.

At this stage, no distinction is made between flows of different services, as this is only an enumeration of feasible paths in topological terms, without taking into account any specifications regarding bandwidth demands for each flow.

In this problem formulation ($\mathcal{P}_0$), the objective is the minimization of the total number of links of the path pair.

$$\min \left\{ \sum_{k \in \mathcal{A}} \left(a_t^{k,1} + a_t^{k,2} \right) \right\} \tag{1}$$

subject to:

$$\sum_{k \in \delta(v^+)} a_t^{k,1} - \sum_{k \in \delta(v^-)} a_t^{k,1} = \begin{cases} 1 & \text{if } v = i \\ -1 & \text{if } v = j \\ 0 & \text{if } v \neq i \text{ and } v \neq j \end{cases} \quad \forall v \in \mathcal{N} \tag{2}$$

$$\sum_{k \in \delta(v^+)} a_t^{k,2} - \sum_{k \in \delta(v^-)} a_t^{k,2} = \begin{cases} 1 & \text{if } v = i \\ -1 & \text{if } v = j \\ 0 & \text{if } v \neq i \text{ and } v \neq j \end{cases} \quad \forall v \in \mathcal{N} \tag{3}$$

$$\sum_{k \in \delta(v^+)} a_t^{k,1} + \sum_{k \in \delta(v^+)} a_t^{k,2} \leq 1 \quad \forall v \in \mathcal{N} \setminus \{i, j\} \tag{4}$$

$$a_t^{(i,j),1} + a_t^{(i,j),2} \leq 1 \quad \text{if there is the direct link } k \equiv (i, j) \tag{5}$$

$$\sum_{k \in \mathcal{A}} a_t^{k,1} \leq D_{ij} \tag{6}$$

$$\sum_{k \in \mathcal{A}} a_t^{k,2} \leq D_{ij} \tag{7}$$

Equations (2) and (3) guarantee that i and j are the origin and the destination nodes (respectively) of the path pair; Equation (4) guarantees that the path pair for flow t is node-disjoint; Equation (5) guarantees that if there is a direct link beginning in the origin node i and destined for the end node j, that link will not be used simultaneously in the active path and the backup path (Equation (4) does not prevent this situation); Equations (6)-(7) impose a limit for the maximal number of links for any path originating in node i and destined for node j.

Equations (6)-(7) are considered, so that the obtained paths are not excessively long, a constraint assumed in most teletraffic engineering approaches. Let $D_{ij}^{\min}$ be the number of links of the shortest path originating in node i and destined for node j. The paths originating in node i and destined for node j will have a maximal number of links given by a logarithmic function: $D_{ij} = D_{ij}^{\min} + 5 * \left\lceil \ln \left(D_{ij}^{\min} \right) \right\rceil$, with $\lceil \cdots \rceil$ representing the ceiling function and $D_{ij}^{\min} > 1$. For $D_{ij}^{\min} = 1$, it is established that $D_{ij} = 5$. Some examples: for $D_{ij}^{\min} = 10$, it is considered that $D_{ij} = 25$; for $D_{ij}^{\min} = 100$, it is considered that $D_{ij} = 125$.

Flows t (and the symmetrical t'), for which this problem has at least one solution, are part of the set $\mathcal{T}_N$.

2.2.2. *Enumeration of the flows with link-disjoint path pairs*

The following optimization problem is solved for each flow $t \equiv (i, j, s) \notin \mathcal{T}_N$, with $i < j$, for enumerating flows with link-disjoint path pairs and a maximal number of links per path. Again it is sufficient to solve the problem for flows $t \equiv (i, j, s)$, with $i < j$, and no distinction is made between flows of different services.

In this problem formulation ($\mathcal{P}_1$), the objective is also the minimization of the total number of links of the path pair.

$$\min \left\{ \sum_{k \in \mathcal{A}} \left(a_t^{k,1} + a_t^{k,2} \right) \right\} \tag{8}$$

$$\text{subject to:} \qquad (2) - (3) \text{ and } (5) - (7)$$

$$a_t^{k,1} + a_t^{k,2} \leq 1 \quad \forall k \in \mathcal{A} \tag{9}$$

Equation (9) guarantees that the path pair for flow t is link-disjoint. Flows t (and the symmetrical t'), for which this problem has at least one solution, are part of the set $\mathcal{T}_A$.

2.2.3. *Enumeration of feasible disjoint path pairs for each flow*

After the identification of all the flows belonging to $\mathcal{T}_N$ and to $\mathcal{T}_A$, the K-shortest path algorithm MPS (Martins *et al.* 1999, Gomes *et al.* 2001), will be used to aid in the enumeration of feasible node-disjoint path pairs for each flow $t \in \mathcal{T}_N$ and feasible link-disjoint path pairs for each flow $t \in \mathcal{T}_A$. It is sufficient to consider the flows $t \equiv (i, j, s)$, with $i < j$.

With the MPS algorithm, $K \geq 1$ shortest paths can be enumerated from a specific origin node to a specific destination node, ordered according to an ascending cost. The cost function of each link is additive, so the cost of a path is the sum of the cost of the links that are part of that path. In this case, the link cost will be 1, so that the cost of a path will be equivalent to the total number of links that constitute the path. Hence, for each flow $t \in \{\mathcal{T}_N \cup \mathcal{T}_A\}$, a set of paths in ascending order of the number of links is obtained. For each of these paths, a modified network is considered, where all the intermediate nodes (links) of the original path are removed and a new set of paths is obtained in this modified network for that flow $t \in \mathcal{T}_N\,(\mathcal{T}_A)$. The original path and each of the obtained paths are node(link)-disjoint, as desired for protection purposes. An example of the application of this procedure for enumerating feasible disjoint path pairs for a flow, in an example network topology, is provided as supplementary material.

As mentioned earlier, a maximal number of links was imposed for each origin-destination pair, so that the obtained paths are not excessively long. This limitation will naturally impose a maximal value for N_t, which represents the number of node(link)-disjoint path pairs that will be considered for each flow $t \in \mathcal{T}_N\,(\mathcal{T}_A)$, based on the topological features of the network and the value of D_{ij}. The possibility of admitting longer paths will increase the number of feasible paths. Notice that no limitations are imposed on N_t based on the class of service s to which the flow t belongs.

For the maximal value for L_t, it is considered that $L_t^{\max} = \min\{N_L, N_t\}$, where N_L is the maximum number of established path pairs, so as to guarantee that the space of the problem solutions does not become too vast.

In the considered algorithm for enumeration of disjoint path pairs, the first shortest path is taken into account and a maximum of $L_t^{\max}$ paths is sought in the modified network. If a total of $L_t^{\max}$ paths is found, then there is a total of $L_t^{\max}$ disjoint path pairs and the search ends; otherwise, the second shortest path is taken into account and a set of paths is sought in the modified network, to try and find more disjoint path pairs; the search continues until $L_t = L_t^{\max}$. If there are some node-disjoint path pairs for a specific origin-destination pair, but not a total of $L_t^{\max}$ path pairs, then it is necessary to try to find some link-disjoint path pairs for that specific origin-destination pair, so that the total of $L_t = L_t^{\max}$ disjoint path pairs is achieved.

With the information on the links of each path pair, a binary value matrix may be obtained: $a_{t,l}^{k,m} = 1$ if link $k \in p_{t,l}^m$ and 0 otherwise, with $m = 1$ or 2; $l = 1, \cdots, L_t; k \in \mathcal{A}; t \in \{\mathcal{T}_N \cup \mathcal{T}_A\}$. This matrix has dimension $|\mathcal{A}| \times (|\mathcal{T}_N| + |\mathcal{T}_A|) \times L_t^{\max}$.

2.3. *Routing optimization model*

The formulated routing optimization problem is biobjective. The two considered objective functions are conflicting functions and the obtained solutions for the routing problem are compromise solutions chosen in the set of non-dominated solutions found by the proposed exact resolution algorithm.

For each network link $k \in \mathcal{A}$, an additive cost per unit of bandwidth, c_k, is considered. This cost is formulated in terms of the link length l_k [km], which is associated with propagation delays, and the inverse of the link capacity u_k [Gbps], which is associated with a scaling effect and a possible decrease in transmission times. Let

$$c_k = \frac{\alpha}{u_k^*} + \beta l_k^* \tag{10}$$

with $\alpha, \beta > 0$. Notice that

$$\frac{1}{u_k^*} = \frac{\frac{1}{u_k} - \min_{\kappa\in\mathcal{A}} \frac{1}{u_\kappa}}{\max_{\kappa\in\mathcal{A}} \frac{1}{u_\kappa} - \min_{\kappa\in\mathcal{A}} \frac{1}{u_\kappa}} \quad \text{and} \quad l_k^* = \frac{l_k - \min_{\kappa\in\mathcal{A}} l_\kappa}{\max_{\kappa\in\mathcal{A}} l_\kappa - \min_{\kappa\in\mathcal{A}} l_\kappa}$$

are normalized versions of the parameters. It is assumed that $\max_{\kappa\in\mathcal{A}} u_\kappa \neq \min_{\kappa\in\mathcal{A}} u_\kappa$. On the contrary, if all the links in the network have the same capacity, then this parameter affects all the links in the same way and does not need to be included in the link cost, leading to $c_k = \alpha + \beta l_k^*$. Likewise, it is assumed that $\max_{\kappa\in\mathcal{A}} l_\kappa \neq \min_{\kappa\in\mathcal{A}} l_\kappa$. On the contrary, if all the links have practically the same length, then this parameter affects all the links in the same way and does not need to be included in the link cost, leading to $c_k = \frac{\alpha}{u_k^*} + \beta$. If both situations occur, then $c_k = \alpha + \beta$, which is actually independent of the link k.

Let the binary decision variables $y_{t,l} = 1$ if the l-th path pair $\left(p_{t,l}^1, p_{t,l}^2\right)$ is the considered solution for routing the flow t, and $y_{t,l} = 0$ otherwise. The decision space $Y = \{y_{t,l}\}$ has dimension $\mathcal{D}_Y = \prod_{t\in\{\mathcal{T}_N\cup\mathcal{T}_A\}} L_t \leq \left(\max_{t\in\{\mathcal{T}_N\cup\mathcal{T}_A\}} \{L_t\}\right)^{|\mathcal{T}_N|+|\mathcal{T}_A|} \leq \left(\max_{t\in\{\mathcal{T}_N\cup\mathcal{T}_A\}} \{L_t\}\right)^{n(n-1)|\mathcal{S}|}$.

The first objective function to be defined is the minimization of the total cost of carrying the bandwidth of all the flows offered to all the feasible path pairs:

$$\min F_1 = \sum_{t\in\{\mathcal{T}_N\cup\mathcal{T}_A\}} \left[\sum_{l=1}^{L_t} \sum_{k\in\mathcal{A}} \left(\left(a_{t,l}^{k,1} + a_{t,l}^{k,2}\right) c_k \xi_{ks}\right) y_{t,l}\right] d_t \tag{11}$$

$$\text{with:} \qquad \sum_{l=1}^{L_t} y_{t,l} = 1 \quad \forall t \in \{\mathcal{T}_N \cup \mathcal{T}_A\} \tag{12}$$

$$y_{t,l} - y_{t',l} = 0 \quad \forall t \in \{\mathcal{T}_N \cup \mathcal{T}_A\}\,; l = 1, \cdots, L_t; i < j \tag{13}$$

$$f_k = \sum_{t \in \{\mathcal{T}_N \cup \mathcal{T}_A\}} \left[\sum_{l=1}^{L_t} \left(a_{t,l}^{k,1} + a_{t,l}^{k,2} \right) y_{t,l} \right] \xi_{ks} d_t \quad \forall k \in \mathcal{A} \tag{14}$$

$$f_k \leq u_k \quad \forall k \in \mathcal{A} \tag{15}$$

The constraint (12) guarantees that only one path pair is used for each flow; Equation (13) is a symmetry constraint, which guarantees that if the l-th path pair $\left(p_{t,l}^1, p_{t,l}^2\right)$ is used for flow t, then the topologically symmetrical path pair $\left(p_{t',l}^1, p_{t',l}^2\right)$ is used for flow t'; Equation (14) is the calculation of the bandwidth f_k carried in each link $k \in \mathcal{A}$ and Equation (15) guarantees that f_k does not exceed the capacity of the link $k \in \mathcal{A}$.

The second objective function in this model is the minimization of the total load cost in the network links. This load cost for link k, ϕ_k, is associated with the effect of carrying the bandwidth of all the flows offered to all the feasible paths which use that link. By considering the network total load cost function, an optimally balanced distribution of the load in the network may be achieved. Therefore the over-utilization of the more loaded links may be avoided, as sending more packets over a link is increasingly penalized as the utilization of the link increases. This approach favours the possibility of the network accepting more traffic in the future. This type of function has been used in earlier studies on routing models in MPLS (Erbas and Erbas 2003, Craveirinha *et al.* 2007, 2013) and was originally proposed in (Fortz and Thorup 2000, 2002).

The load cost function for each link, ϕ_k, is piece-wise linear and it is defined for each link $k \in \mathcal{A}$, based on the total load (amount of flow) f_k carried in the link and the link capacity u_k. An example of this type of function is in (Erbas and Erbas 2003, Fig.1).

Hence, the second objective function is the minimization of the network total load cost:

$$\min F_2 = \sum_{k \in \mathcal{A}} \phi_k \tag{16}$$

$$\text{with:} \qquad (12) - (15)$$

$$\phi_k \geq f_k, \quad \forall k \in \mathcal{A} \tag{17}$$

$$\phi_k \geq 2 f_k - 0.5 u_k, \quad \forall k \in \mathcal{A} \tag{18}$$

$$\phi_k \geq 5 f_k - 2.3 u_k, \quad \forall k \in \mathcal{A} \tag{19}$$

$$\phi_k \geq 15 f_k - 9.3 u_k, \quad \forall k \in \mathcal{A} \tag{20}$$

$$\phi_k \geq 60 f_k - 45.3 u_k, \quad \forall k \in \mathcal{A} \tag{21}$$

$$\phi_k \geq 300 f_k - 261.3 u_k, \quad \forall k \in \mathcal{A} \tag{22}$$

where (17)-(22) define the piece-wise linear function to be considered.

The addressed biobjective routing optimization problem with path protection may be formulated as problem $\mathcal{P}_2$:

$\min\{F_1, F_2\}$	
subject to:	(12)-(15), (17)-(22)
	Paths for flows $t \in \{\mathcal{T}_N \cup \mathcal{T}_A\}$ in $\mathcal{L}_t$

Although the structure of this problem is of the same type (a bicriteria routing optimization problem) as the one in (Girão-Silva *et al.* 2015), notice that the problem in itself is quite different. In (Girão-Silva *et al.* 2015), the possibility of traffic splitting was taken into account and there was no path protection schemes, leading to a different problem in terms of the formulation of the objective function F_1 and of the problem constraints.

3. Resolution method

In the considered resolution method, after the enumeration of feasible path pairs is performed, an exact resolution method based on an adaptation of the classical constraint method is applied in a second stage to solve the formulated problem. In a strict sense, the resolution method viewed as a whole (i.e. considering both stages of calculation of candidate path pairs and actual choice of adequate path pairs for each flow) is not exact because there is the possibility that some path pairs may not be considered as results of the first stage. However, the resolution approach would be exact if all the feasible path pairs were considered as inputs in the second stage. Of course this might lead, in some cases, to an excessive computational burden or even to computational intractability in the second stage.

Concerning the second stage of the resolution approach, recall that in the classical constraint method (Cohon 1978), one of the objective functions is kept, while all the others are transformed in auxiliary constraints. Therefore, an auxiliary single objective problem is formulated and it is easily solvable by conventional methods. As explained in (Cohon 1978), the optimal solution of this single objective problem is a non-dominated solution of the original multiobjective problem. Different single objective problems are obtained when different bounds are imposed on the constrained objectives. These bounds must be chosen, so as to guarantee that there is an optimal solution to each single objective problem, corresponding to a non-dominated solution of the original problem. This exact method is able to calculate all non-dominated solutions, either supported or unsupported.

In order to clarify the idea behind the classical constraint method, consider Figure 1 . In this figure the extreme solutions of the Pareto front are shown, where $X \equiv (F_1^{\min}, F_2^{\max})$ and $Y \equiv (F_1^{\max}, F_2^{\min})$. These points delimit the original feasible region. A single objective problem is considered, the minimization of the objective function F_2, and a constraint for the other objective function, $F_1 \leq F_{1\text{lim}}$, is formulated. With this constraint, a new feasible region (dashed portion in Figure 1) is established where the optimization of F_2 is sought. By successively changing the limit $F_{1\text{lim}}$, a set of non-dominated solutions for a biobjective problem may be obtained, as shown in the example of Figure 2.

The considered resolution method is presented in (Girão-Silva *et al.* 2015). The classical constraint method is at the core of the proposed resolution method, but special features were added that are worth mentioning: i) finding weakly non-dominated solutions is avoided (Mavrotas 2009), both when the pay-off table for the objective functions F_1 and F_2 is being obtained and when the single objective problem is being solved; ii) after finding solutions in the bidimensional objective function space, preference thresholds for the two objective functions are defined, that will allow for the definition of preference regions in this space; the region considered to be the most preferable will be further explored, so that more non-dominated solutions can be found in that particular region; iii)

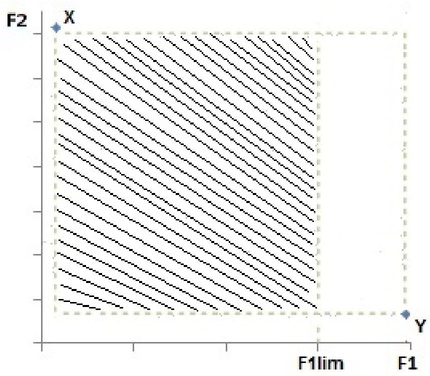

Figure 1. Example of the new feasible region in the context of the classical constraint method.

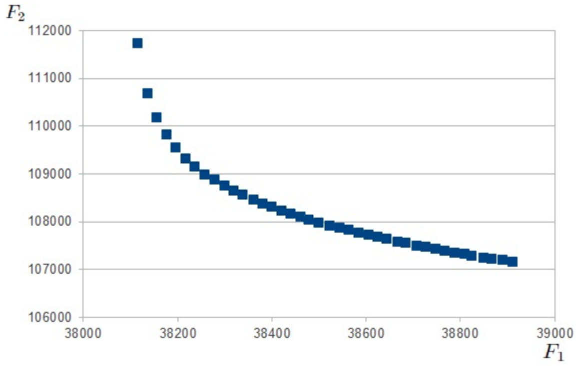

Figure 2. Example of the Pareto front obtained for a network with $|\mathcal{N}| = 30$ nodes and $\eta = 4.0$, with $\mathcal{K} = 22\%$.

considering all the obtained non-dominated solutions in the most preferable region in the bidimensional objective function space, a specific solution will be chosen; in particular, it will be the one which minimizes a weighted Chebyshev distance to a reference point.

The formalization of the algorithm used for solving problem $\mathcal{P}_2$, is available as supplementary material.

4. Experimental study

An experimental study for evaluation of the proposed method was carried out in a set of randomly generated MPLS networks for which the topologies were obtained with the gt-itm software (GT-ITM 2000), for a certain number of nodes and average node degree. In these networks, the capacities of the links and the offered traffic matrix have to be specified. Therefore, a dimensioning algorithm had to be devised and implemented for calculating the capacities of the links, once the network structures and the offered traffic matrices, to be used in the experimental study, are known.

For each network topology, 5 different instances of the problem are considered, each instance corresponding to an offered traffic matrix (the demand volumes are uniformly generated). The link capacities for each instance were dimensioned according to the offered traffic matrix so that the problem "actually made sense" and could be solved.

In the next sub-sections, the features of the networks and of the considered services are described. Afterwards, the performed tests are described and the network performance measures are presented. Finally, the obtained results are analyzed.

4.1. *Service and network description*

Different types of services were considered, with $\mathcal{S}$ denoting the set of offered services. A total of $|\mathcal{S}| = 4$ service profiles were taken into account, with different bandwidth values: $\ell_1 = 5$ Mbps ($s = 1$), $\ell_2 = 50$ Mbps ($s = 2$), $\ell_3 = 500$ Mbps ($s = 3$) and $\ell_4 = 5$ Gbps ($s = 4$). For each service s and each origin-destination pair (i, j), a specific value of demand is considered, given by $d_t = d_{(i,j,s)}$, that is a multiple of the bandwidth value for that specific service.

A set of randomly generated MPLS network topologies obtained with the gt-itm software is considered, given a particular number of nodes $|\mathcal{N}|$ and a particular average node degree η. This software allows for the creation of realistic structures of Internet-type networks (Doar and Leslie 1993). The Doar-Leslie model (Doar and Leslie 1993) has been considered for the probability distribution of the existence of links, calibrated so as to guarantee a value for the average node degree η, considering the specific number of nodes $|\mathcal{N}|$. This software generates a network topology where the nodes are connected by a pair of directed links. The maximal value of the link length is $l_{\max} = 40$ km, a typical value for metropolitan networks.

In particular, a wide variety of experiments have been performed in networks with $|\mathcal{N}| = 14$ nodes and $\eta = 2.7$ and $\eta = 4.0$, as well as in networks with $|\mathcal{N}| = 30$ nodes and $\eta = 2.7$ and $\eta = 4.0$. Some experiments were also performed in larger networks, in particular, networks with $|\mathcal{N}| = 50$, $|\mathcal{N}| = 100$ and $|\mathcal{N}| = 150$ nodes.

The traffic matrices values were randomly generated according to a uniform distribution. Therefore, for a flow $t \equiv (i, j, s)$ (associated with service s, with origin in node i and destined for node j), $d_t = \nu \ell_s$ is calculated, with $\nu \in \mathbb{N}$ uniformly distributed between 0 and 9. Given the information on the traffic matrices values for all the services and all the origin-destination pairs, the total traffic volume originating in any node i and destined for any node j, $\ell(i,j) = \sum_{s\in\mathcal{S}} d_{(i,j,s)}$, is calculated.

The link capacities in the network are calculated by a simple dimensioning procedure: i) at first, all links are assumed to have null capacity; ii) for each origin-destination pair (i, j), the shortest path in the network (in terms of the number of links) is considered; iii) assuming this is the path used for that origin-destination pair (i, j), the links in that path have their capacity increased by a value of $\frac{\ell(i,j)}{\mathcal{K}}$, where $\mathcal{K} \in [0.0; 1.0]$. The value $\mathcal{K}$ represents a percentage value and in this way there is a guarantee that only a percentage of the link capacity will be used. For instance, if $\mathcal{K} = 0.2 = 20\%$ and if the total required bandwidth of a link is 5 Gbps, then the total link capacity will be $\frac{5}{0.2} = 25$ Gbps, that is, only a fraction given by $\mathcal{K}$ will be used to carry the required bandwidth.

After this dimensioning operation, all the link capacities will be rounded according to a ceiling function, so that all the links have a capacity that is an integer value in Gbps. For all pairs of directed links, the final value of link capacity is the highest of the pair, i.e. if link k has a capacity $\overline{u}_k$ and link k' has a capacity $\overline{u}_{k'}$, then the final values will be $u_k = u_{k'} = \max\{\overline{u}_k; \overline{u}_{k'}\}$.

4.2. *Network performance measures*

Numerous experiments were carried out with the considered topologies. In particular, 5 instances of offered traffic matrices (and corresponding link capacities) were obtained, so that a total of 5 different experiments were performed for each network topology.

The values of some network performance measures were calculated, so that the quality of the routing solution obtained with the method can be evaluated. Some of these performance parameters (involving the link utilization) are 'standard' measures of network performance and they are often used in the evaluation of routing models (see (Srivastava *et al.* 2005) for details on the use of these measures in a context of single criterion optimization approaches).

For each problem instance (i.e. a network structure with a traffic matrix), 3 different solutions are obtained: S_1, the solution obtained when the objective function F_1 is minimized; S_2, the solution obtained when the objective function F_2 is minimized; S_M, the solution obtained when the biobjective routing method is used.

The following performance measures were defined:

- total fraction of used capacity: $FUC = \frac{\sum_{k\in\mathcal{A}} f_k}{\sum_{k\in\mathcal{A}} u_k}$;
- sum of the link utilizations: $SLU = \sum_{k\in\mathcal{A}} \frac{f_k}{u_k}$;
- average link utilization: $ALU = \frac{SLU}{|\mathcal{A}|} = \frac{1}{|\mathcal{A}|}\sum_{k\in\mathcal{A}} \frac{f_k}{u_k}$;
- maximal link utilization: $MLU = \max_{k\in\mathcal{A}} \left\{\frac{f_k}{u_k}\right\}$;
- relative variation with respect to the marginal optima: $RV_\varrho = \left|\frac{F_\varrho^{\text{sol}} - F_\varrho^{\text{opt}}}{F_\varrho^{\text{opt}}}\right|$ (with $\varrho = 1, 2$), where F_ϱ^{sol} is the value of F_ϱ calculated for a specific biobjective solution and F_ϱ^{opt} is the optimal value of F_ϱ for the same problem.
- time of execution of the algorithms in a computer with a 3.07 GHz clock and 6 GB of RAM, with the Linux operating system.

4.3. *Description of the performed experiments and obtained results*

The cost of each link is given by c_k (see Equation (10)), with $\alpha = 0.1$ and $\beta = 1-\alpha = 0.9$, which means more weight is given to the link length than to the inverse of the link capacity.

It is assumed that the maximal number of path pairs considered for routing each flow is $N_L = 10$. For the execution of the resolution algorithm (see details in the supplementary material), it is considered that $\gamma = 1E^{-6}$ and $\Omega = 10$. All the overbooking factors were made equal to 1.

4.3.1. *Experiments with $\mathcal{K} = 20\%$ and networks of 14 and 30 nodes*

A total of 4 different topologies was considered for each pair $(|\mathcal{N}|; \eta)$. For each topology, 5 instances were considered, i.e. 5 different traffic matrices and capacity values for the network links (obtained as explained in section 4.1, with $\mathcal{K} = 20\%$). For these instances, exact results were obtained, which allow for the calculation of the minimum, the average and the maximum of each network metric (objective functions and performance measures) for each topology. In order to present the results in a more compact way, the obtained values of the minimum, the average and the maximum value of each parameter for each of the 4 different topologies for each pair $(|\mathcal{N}|; \eta)$ were considered and their mean value and their variation range were estimated. Considering the value of a performance measure for the z-th topology as $\mathcal{M}_z$, with $z = 0, \cdots, 3$, the estimate of the mean is $\hat{\mathcal{M}} = \frac{1}{4}\sum_{z=0}^{3} \mathcal{M}_z$ and the estimate of the variation range is $\frac{3.182}{\sqrt{4}}\hat{\sigma}(\hat{\mathcal{M}})$, calculated as a 95% confidence interval with a t-Student bi-lateral distribution, where $\hat{\sigma}(\hat{\mathcal{M}}) = \sqrt{\frac{\sum_{z=0}^{3}(\mathcal{M}_z - \hat{\mathcal{M}})^2}{3}}$ is an estimate of a standard deviation for $\mathcal{M}$ (Banks and Carson 1984). Further information on the display of the obtained results is available as supplementary material.

The results for the networks with $|\mathcal{N}| = 30$ nodes are in Tables 1-2. In Figures 3-4, the average results obtained for some of the performance measures in networks with $|\mathcal{N}| = 30$ nodes and average node degree $\eta = 2.7$, with $\mathcal{K} = 20\%$ – see Table 1 – are displayed. These results are representative of the general trend of the results. Further results (for networks with $|\mathcal{N}| = 14$ nodes) are available as supplementary material. The displayed results were obtained using the CPLEX 12.6 software (CPLEX Software 2013)

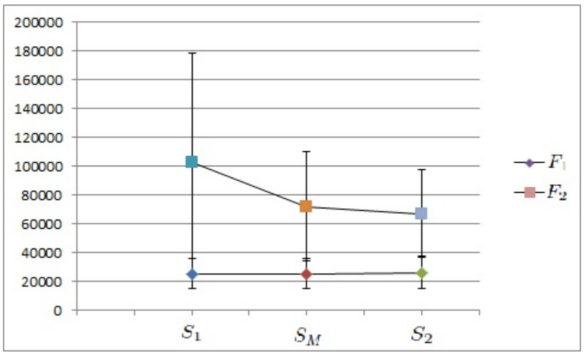

Figure 3. Results for the objective functions F_1 and F_2.

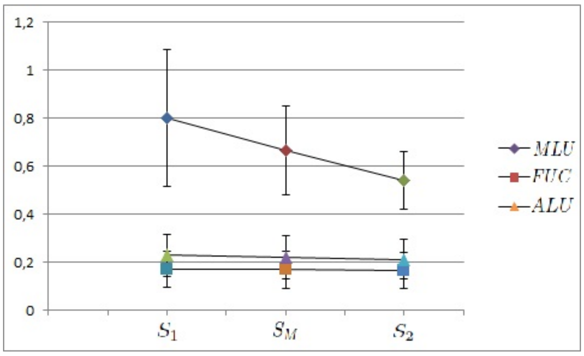

Figure 4. Results for FUC, MLU and ALU performance measures.

for resolution of the single objective integer problems in steps 2 and 3 of the algorithm, for which the formalization is available as supplementary material. For the resolution of an integer problem, the software begins with a pre-solver that will help in reducing the problem size, followed by a branch-and-bound method to actually solve the problem.

Table 1. Results for the networks with $|\mathcal{N}| = 30$ nodes and average node degree $\eta = 2.7$, with $\mathcal{K} = 20\%$.

Perf. measure	Sol.	Minimum	Average	Maximum
F_1	S_1	23963.18 ± 10070.87	25292.02 ± 10385.88	26253.70 ± 10454.24
	S_2	24889.62 ± 11220.42	26158.37 ± 11168.00	27130.35 ± 11158.00
	S_M	24131.70 ± 10315.96	25431.38 ± 10565.15	26389.74 ± 10627.10
F_2	S_1	80902.25 ± 55266.66	102318.39 ± 76029.49	134589.90 ± 118217.41
	S_2	63032.01 ± 28958.36	66952.52 ± 30371.27	69744.72 ± 31962.29
	S_M	66911.49 ± 34655.87	72198.20 ± 37570.05	77237.43 ± 41634.67
RV_1	S_2	0.0280 ± 0.0184	0.0327 ± 0.0199	0.0396 ± 0.0272
	S_M	0.0039 ± 0.0033	0.0050 ± 0.0045	0.0066 ± 0.0062
RV_2	S_1	0.1697 ± 0.2024	0.4755 ± 0.6051	0.8700 ± 1.1764
	S_M	0.0384 ± 0.0417	0.0694 ± 0.0735	0.1116 ± 0.1264
FUC	S_1	0.1632 ± 0.0645	0.1704 ± 0.0754	0.1771 ± 0.0830
	S_2	0.1575 ± 0.0634	0.1645 ± 0.0748	0.1705 ± 0.0825
	S_M	0.1613 ± 0.0658	0.1685 ± 0.0772	0.1751 ± 0.0854
MLU	S_1	0.7098 ± 0.2487	0.8012 ± 0.2851	0.8783 ± 0.3121
	S_2	0.4968 ± 0.1196	0.5414 ± 0.1200	0.5915 ± 0.1688
	S_M	0.6109 ± 0.1727	0.6672 ± 0.1851	0.7437 ± 0.1744
SLU	S_1	18.2561 ± 7.8688	18.8407 ± 8.1042	19.3259 ± 8.0256
	S_2	16.9705 ± 7.1538	17.5202 ± 7.5959	17.9766 ± 7.6917
	S_M	17.6579 ± 7.7915	18.2952 ± 8.1842	18.7946 ± 8.2743
ALU	S_1	0.2213 ± 0.0864	0.2284 ± 0.0892	0.2344 ± 0.0878
	S_2	0.2057 ± 0.0779	0.2125 ± 0.0842	0.2181 ± 0.0853
	S_M	0.2140 ± 0.0857	0.2218 ± 0.0904	0.2278 ± 0.0909
time [s]	S_M	100.06 ± 43.10	109.48 ± 45.10	116.19 ± 46.19

The results validate the initial assumption that the objective functions F_1 and F_2 are

Table 2. Results for the networks with $|\mathcal{N}| = 30$ nodes and average node degree $\eta = 4.0$, with $\mathcal{K} = 20\%$.

Perf. measure	Sol.	Minimum	Average	Maximum
F_1	S_1	39696.05 ± 6070.44	40745.79 ± 5607.06	41859.32 ± 6417.36
	S_2	41873.43 ± 7045.68	42935.46 ± 6675.52	44242.86 ± 7513.20
	S_M	39902.51 ± 6179.97	40952.12 ± 5681.82	42069.77 ± 6470.45
F_2	S_1	250786.70 ± 192838.11	303026.53 ± 243102.28	369427.37 ± 297080.56
	S_2	108959.37 ± 16280.41	111832.75 ± 15077.86	114924.78 ± 15225.33
	S_M	123782.63 ± 25939.25	128348.02 ± 28030.64	132746.85 ± 28910.35
RV_1	S_2	0.0455 ± 0.0249	0.0531 ± 0.0338	0.0595 ± 0.0385
	S_M	0.0040 ± 0.0011	0.0050 ± 0.0020	0.0067 ± 0.0036
RV_2	S_1	1.2083 ± 1.4817	1.6758 ± 1.8949	2.3121 ± 2.3497
	S_M	0.1141 ± 0.1020	0.1452 ± 0.1326	0.1766 ± 0.1662
FUC	S_1	0.3899 ± 0.0593	0.3948 ± 0.0586	0.4041 ± 0.0593
	S_2	0.3723 ± 0.0495	0.3770 ± 0.0491	0.3858 ± 0.0497
	S_M	0.3837 ± 0.0520	0.3881 ± 0.0516	0.3964 ± 0.0523
MLU	S_1	0.9567 ± 0.0885	0.9878 ± 0.0249	0.9995 ± 0.0017
	S_2	0.5608 ± 0.0772	0.5862 ± 0.1093	0.6272 ± 0.1585
	S_M	0.7422 ± 0.0672	0.7667 ± 0.0561	0.7944 ± 0.0355
SLU	S_1	53.0301 ± 10.2853	53.9988 ± 10.2154	55.0191 ± 10.7470
	S_2	49.9922 ± 7.1888	50.7573 ± 7.2701	51.7126 ± 7.5601
	S_M	51.8068 ± 8.6422	52.6563 ± 8.8547	53.5318 ± 9.3385
ALU	S_1	0.4153 ± 0.0624	0.4229 ± 0.0606	0.4308 ± 0.0624
	S_2	0.3918 ± 0.0408	0.3978 ± 0.0405	0.4052 ± 0.0406
	S_M	0.4059 ± 0.0507	0.4125 ± 0.0505	0.4193 ± 0.0522
time [s]	S_M	802.19 ± 1514.66	1272.40 ± 2431.92	1739.19 ± 3105.42

conflicting, as the decrease of one of them leads to an increase in the value of the other objective function. In fact, when a more balanced distribution of the traffic is achieved, which is accompanied by a decrease in the value of F_2, the routing cost (expressed by F_1) increases. This trend is also noticeable in Figure 3. The variation range in the values of F_2 is higher than in the values of F_1, especially for the S_1 solutions (i.e. the solutions obtained when F_1 is optimized individually). The clear conflicting nature between F_1 and F_2 has already been noted in earlier studies in single criterion network flow models, as in (Srivastava *et al.* 2005).

The minimization of F_1 results in the lowest total cost of carrying the bandwidth of all the flows and it tends to originate a lower number of links being used, but the links are used in a more intensive way, leading to a worsening in the values of the network performance measures related to the link utilization. This trend is confirmed, as when only the objective function F_1 is optimized (solutions S_1) an increase in the utilization of the links in the majority of the tested networks is noticeable. In fact, the values of FUC, SLU, MLU and ALU tend to be higher than when only the objective function F_2 is optimized (solutions S_2) or when the biobjective problem is considered (solutions S_M). This trend is also noticeable in Figure 4, in particular for the value of MLU, which is substantially higher for the S_1 solutions than for the other solutions. The compromise solutions S_M provide an intermediate value for these performance measures. Although a lower value for these performance measures would be preferable, bear in mind that the total cost of carrying the bandwidth of all the flows also has to be taken into account and therefore, the adequacy of the formulated resolution method becomes more evident.

For solutions S_2 (i.e. resulting from the optimization of F_2), the cost of carrying the bandwidth of all the flows increases, which is visible in the value of F_1. The global utilization of the links tends to be lower, which is expectable when F_2 is minimized. The decrease in the utilization of the links can be confirmed both by the lower value of F_2 and by the lower values of the relevant performance measures (FUC, SLU, MLU and ALU).

The solutions S_M lead to compromise values for the performance measures, as expected. The results of these experiments confirm the potential advantages of using a multiobjective optimization model, rather than a single objective one to solve this routing problem. The advantages of using two objectives become clear, as the obtained solution represents a compromise between the cost of carrying the bandwidth and the global effect of the utilization of the links.

As for the computing times, they increase with the increase in the problem dimension, as expected. This may be the most important limitation of the developed exact resolution method, as problems with large networks (a large number of nodes and/or links) may become computationally untractable.

As mentioned earlier, the networks were dimensioned so that only $\mathcal{K} = 20\%$ of the link capacity was used if the shortest paths were used to carry each flow. A value of FUC close to $2\mathcal{K} = 40\%$ would be expected, as a pair of paths is used by each flow in the routing model. However, for some networks, the value of FUC is lower (see for example Table 1). This is due to the fact that in these networks there are many nodes from and to which no disjoint path pairs can be found. Although the traffic from and to these nodes is not considered in the final problem to be solved, the networks were dimensioned having that traffic into account.

4.3.2. *Experiments with $\mathcal{K} = 22\%$ and networks of 30 nodes*

The running time of the algorithm is expected to increase if the capacity of the links decreases, as it becomes more difficult to "accommodate" the offered traffic, and it takes longer to obtain an adequate solution. In order to check this effect, the routing algorithm was run for the networks with $(|\mathcal{N}|; \eta) = (30; 4.0)$, but this time with $\mathcal{K} = 0.22 = 22\%$. This will result in a lower link capacity value. For instance, if $\mathcal{K} = 0.22 = 22\%$ and if the total required bandwidth of a link is 5 Gbps, then the total link capacity will be $\frac{5}{0.22} = 22.7 \approx 23$ Gbps.

The aggregated results are in Table 3 and can be compared with those in Table 2. The values of F_1 and F_2 and also RV_1 and RV_2 cannot be directly compared, but the values of the performance measures and the execution time can be compared. As expected, all the values regarding the utilization of the links are higher when $\mathcal{K} = 22\%$ as the link capacities obtained in the dimensioning process are smaller. As the network is more heavily loaded, it is more difficult to find adequate solutions and therefore, the execution times are much higher: on average, the execution time is 1272 s when $\mathcal{K} = 20\%$ and 7044 s when $\mathcal{K} = 22\%$ (an increase of 5.5 times). This experiment shows that the execution times may become prohibitive if the networks are heavily loaded. Along with the fact that this method may become computationally untractable for large networks, this may be other important limitation to take into account.

These results (as those in sub-section 4.3.1) were obtained with a relative gap tolerance parameter of 1E-4 in the auxiliary CPLEX procedure (default relative gap tolerance). This is the value for which the obtained solution may be considered "exact" in numerical terms. The relative gap tolerance parameter value (which will be designated in this text by tolerance) in an integer problem can be defined by the user and is a "relative tolerance on the gap between the best integer objective and the objective of the best

Table 3. Results for the networks with $|\mathcal{N}| = 30$ nodes and average node degree $\eta = 4.0$, with $\mathcal{K} = 22\%$.

Perf. measure	Sol.	Minimum	Average	Maximum
F_1	S_1	39294.87 ± 9593.17	40524.94 ± 9825.15	41846.04 ± 9723.91
	S_2	41574.70 ± 11304.39	43245.61 ± 12162.60	44698.01 ± 11362.63
	S_M	39490.33 ± 9579.67	40766.19 ± 9926.63	42114.16 ± 9826.91
F_2	S_1	265372.07 ± 177373.54	318070.83 ± 221067.81	367946.23 ± 291331.52
	S_2	107501.75 ± 25585.32	111674.95 ± 28784.48	116541.51 ± 29491.65
	S_M	121814.76 ± 32747.89	130884.70 ± 43177.98	139922.08 ± 49271.35
RV_1	S_2	0.0567 ± 0.0436	0.0661 ± 0.0465	0.0763 ± 0.0663
	S_M	0.0044 ± 0.0016	0.0059 ± 0.0026	0.0078 ± 0.0052
RV_2	S_1	1.3599 ± 0.8660	1.8105 ± 1.3541	2.1916 ± 1.8889
	S_M	0.1213 ± 0.0603	0.1691 ± 0.1169	0.2068 ± 0.1867
FUC	S_1	0.4071 ± 0.0794	0.4145 ± 0.0690	0.4203 ± 0.0627
	S_2	0.3922 ± 0.0756	0.3996 ± 0.0662	0.4056 ± 0.0602
	S_M	0.4016 ± 0.0713	0.4092 ± 0.0607	0.4148 ± 0.0556
MLU	S_1	0.9882 ± 0.0220	0.9976 ± 0.0043	1.0000 ± 0.0000
	S_2	0.5691 ± 0.1332	0.6509 ± 0.2335	0.7383 ± 0.3999
	S_M	0.7374 ± 0.0176	0.7930 ± 0.0561	0.8384 ± 0.1649
SLU	S_1	54.5690 ± 5.3589	55.8247 ± 4.5981	56.6204 ± 4.1083
	S_2	51.9325 ± 4.5880	52.9933 ± 3.5357	53.7215 ± 2.9536
	S_M	53.3745 ± 4.3722	54.6747 ± 3.5403	55.5940 ± 3.1758
ALU	S_1	0.4358 ± 0.0696	0.4458 ± 0.0633	0.4521 ± 0.0593
	S_2	0.4148 ± 0.0618	0.4232 ± 0.0543	0.4290 ± 0.0498
	S_M	0.4263 ± 0.0612	0.4366 ± 0.0540	0.4439 ± 0.0505
time [s]	S_M	1992.58 ± 5244.77	7043.72 ± 22730.60	11764.72 ± 38130.16

node[1] remaining" (CPLEX Documentation 2013), i.e. if the best and the second best solution found thus far differ by a relative value not greater than this tolerance value, the algorithm stops and the best solution found thus far is presented. The tolerance parameter may be tuned, to achieve a compromise between the quality of the solution and the execution time: if the tolerance parameter is high, then the found solution is expected to be more distant to the exact solution but it can be found faster as the algorithm halts sooner.

In these experiments, the "exact" value (obtained with a tolerance default value of 1E-4) of the solution is known and it can be compared to the solutions obtained with other tolerance values. Experiments were conducted with tolerance values larger than the default value (2E-4, 5E-4, 1E-3, 2E-3 and 5E-3). In Table 4, the aggregated results for the topologies with $(|\mathcal{N}|; \eta) = (30; 4.0)$ and tolerance 5E-3, are presented. For the other values of tolerance, the results follow a pattern consistent with this one and so they have been omitted. In parenthesis, there is the relative error between the sub-optimal solution and the exact one for the values of F_1 and F_2. For example, in Table 4 the average value for $F_1(S_2)$ is 43010.10, which represents a relative error of $|\frac{43010.10-43245.61}{43245.61}| = 0.54\%$ since the corresponding value in Table 3 (table of "exact" values) is 43245.61.

As expected, the relative error between the sub-optimal solution and the "exact" one is of the same order or even inferior to the established value for the tolerance. For instance, when the compromise solution S_M is considered, the relative error for the values

[1]Note that in this context node refers to an element in the decision tree constructed by the branch-and-bound algorithm used to solve the optimization problem.

Table 4. Results for the networks with $|\mathcal{N}| = 30$ nodes and average node degree $\eta = 4.0$, with $\mathcal{K} = 22\%$ and tolerance of 5E-3.

Perf. measure	Sol.	Minimum	Average	Maximum
F_1	S_1	39344.11 ± 9516.60 (0.13%)	40574.90 ± 9774.88 (0.12%)	41871.41 ± 9707.34 (0.06%)
	S_2	41452.68 ± 10893.43 (0.29%)	43010.10 ± 11583.54 (0.54%)	44444.00 ± 10927.86 (0.57%)
	S_M	39694.60 ± 9587.18 (0.52%)	40941.51 ± 9992.45 (0.43%)	42210.85 ± 9801.45 (0.23%)
F_2	S_1	148853.89 ± 72586.32 (43.91%)	200031.62 ± 146862.40 (37.11%)	281872.54 ± 301337.86 (23.39%)
	S_2	107668.94 ± 25421.32 (0.16%)	111816.39 ± 28949.71 (0.13%)	116678.89 ± 29577.56 (0.12%)
	S_M	115737.96 ± 28935.23 (4.99%)	123880.32 ± 40141.75 (5.35%)	132984.07 ± 49411.32 (4.96%)
RV_1	S_2	0.0531 ± 0.0300	0.0593 ± 0.0380	0.0694 ± 0.0576
	S_M	0.0064 ± 0.0036	0.0090 ± 0.0045	0.0116 ± 0.0061
RV_2	S_1	0.3528 ± 0.4260	0.7646 ± 0.9624	1.4433 ± 2.2775
	S_M	0.0695 ± 0.0381	0.1055 ± 0.0861	0.1600 ± 0.1936
FUC	S_1	0.4062 ± 0.0758	0.4134 ± 0.0674	0.4192 ± 0.0594
	S_2	0.3922 ± 0.0752	0.3997 ± 0.0657	0.4057 ± 0.0598
	S_M	0.3999 ± 0.0681	0.4069 ± 0.0606	0.4129 ± 0.0552
MLU	S_1	0.8618 ± 0.0984	0.9174 ± 0.0715	0.9701 ± 0.0920
	S_2	0.5758 ± 0.1040	0.6578 ± 0.2110	0.7382 ± 0.4002
	S_M	0.7136 ± 0.0461	0.7578 ± 0.1180	0.8118 ± 0.2429
SLU	S_1	54.3500 ± 4.6695	55.4785 ± 4.4983	56.4376 ± 4.3174
	S_2	51.9322 ± 4.4285	53.0002 ± 3.5690	53.8072 ± 2.8343
	S_M	35.1066 ± 73.9962	50.6322 ± 12.9740	55.3339 ± 3.7871
ALU	S_1	0.4341 ± 0.0638	0.4430 ± 0.0620	0.4507 ± 0.0602
	S_2	0.4147 ± 0.0606	0.4232 ± 0.0545	0.4296 ± 0.0489
	S_M	0.4242 ± 0.0576	0.4336 ± 0.0547	0.4418 ± 0.0541
time [s]	S_M	218.32 ± 53.24 (89.04%)	251.80 ± 85.09 (96.43%)	288.44 ± 153.99 (97.55%)

of F_1 and F_2 increases as the tolerance parameter increases, reaching a maximal value of about 5.35% (see Table 4), but that is accompanied by a noticeable reduction in the execution time (going from an average value of 7043.72 s on Table 3 to 251.80 s on Table 4, i.e. only 3.57% of the original time). Therefore, the use of the tolerance parameter (adequately tuned to the problem that is being solved) opens up more possibilities concerning the use of this exact method in larger problems (i.e. problems in networks with higher dimensions) or problems in more heavily loaded networks.

4.3.3. *Experiments with $\mathcal{K} = 20\%$ and networks of more than 30 nodes*

Experiments in larger networks were performed in order to evaluate the behavior of the execution times. Following the observation that the final solutions were close to the "exact" ones when tolerance parameters with values higher than the default were used, experiments were performed for networks with 50, 100 and 150 nodes and average node degree of 2.7 and 4.0, considering $\mathcal{K} = 20\%$ and a tolerance of 5E-3. The average values of the execution times are in Table 5.

The increase in execution time is noticeable when the network dimensions increase. For the same number of nodes and different values of average node degree (i.e. different

Table 5. Average execution times [s] for networks with $|\mathcal{N}| = 50$ to $|\mathcal{N}| = 150$ nodes, with $\mathcal{K} = 20\%$ and tolerance of 5E-3.

		Number of nodes		
		50	100	150
Average	2.7	491.95	2899.87	22157.57
node degree	4.0	938.71	6342.14	–

number of links), the time roughly doubles when η goes from 2.7 to 4.0. When the number of nodes doubles (and for the same value of η), there is an increase of 6 to 7.5 times in the execution time. The increase in the execution time is exponential with the increase in the network dimension.

These results confirm the limitations of this exact method for the resolution of problems in large networks. Even with a high tolerance value, the times are very high for networks of medium size (150 nodes). Nevertheless, the method remains compatible with applications in core and metropolitan type networks of significant dimension, which covers most of the practical applications envisaged for the proposed approach. Notice that a similar model and resolution approach were applied in the context of a research and development contract with Altice Labs (formerly Portugal Telecom Inovação), where results in a different experimental framework were obtained (with the public domain software GLPK) and showed to be consistent with those presented here.

5. Conclusions

In this paper, a biobjective routing model for MPLS networks with different service types and path protection was presented. A network-wide optimization approach based on a network flow modeling approach was considered. The routing problem was formulated as a biobjective integer program. The two objectives were the minimization of the bandwidth transport cost and the minimization of the total load cost in the network links. A set of feasible path pairs is devised for each traffic flow and in this context an exact method, based on the classical constraint method, was developed and applied for solving the formulated problem.

A very extensive experimental study was performed, considering MPLS networks with randomly generated topologies and randomly generated traffic matrices. Some of the results on relevant network performance measures, were presented. Examples are the relative variation of some network performance metrics for the obtained solutions, when compared to the values resulting from the solutions obtained when either F_1 or F_2 was individually optimized.

The obtained results show that the objective functions F_1 and F_2 are conflicting, as expected. When the function F_1 is minimized, there is a tendency to route the traffic using the least cost links that become more heavily loaded. When the function F_2 is minimized, there is a tendency to distribute the traffic in a more balanced way. This means that the carried bandwidth decreases in the least cost links and might increase in other links that were not so heavily loaded. This observation confirms the potential advantages of using these two objectives in this routing model, rather than solving a single objective problem. The minimization of F_2 is associated with a decrease in the values of the network performance parameters related to the link utilization. As for the minimization of F_1, since it implies a more intensive usage of less costly links, the values of the performance parameters worsen. The solution provided by the method is a compromise one and it presents intermediate values for the performance parameters, as expected.

The running time increases as the dimension and/or connectivity of the network increase. In order to obtain more suitable running times and broaden the application environment of the proposed exact resolution method, tolerance parameter values were established, that will lead to finding very good quality sub-optimal solutions rather than an "exact" solution. The use of these tolerance parameter values seems adequate, as only in a small fraction of situations will the relative error between the sub-optimal solution and the "exact" solution be larger than the tolerance by a significant order of magnitude. Therefore, even when a tolerance value is established and sub-optimal solutions are obtained, there are still adequate solutions and quite close to the "exact" solution.

The main limitations of the resolution method only become evident when larger networks are taken into account. Due to a larger number of nodes and/or links, the number of potential solutions highly increases and a combinatorial explosion occurs, preventing the software from finding an "exact" solution in a reasonable time.

Concerning the final choice of a non-dominated solution (in the non-dominated solution set), this is performed in an automated manner by a procedure based on the definition of preference regions in the objective function space. Nonetheless, the resolution method can be adapted to an interactive procedure where a decision maker can consider different non-dominated solutions obtained throughout the execution of the algorithm and choose a solution that will best fit his/her preferences, as the different solutions represent different trade-offs between routing cost and total load cost. The number of solutions that can be obtained can be widened straightforwardly by increasing the parameter Ω in the algorithm. Also note that smaller regions in the objective function space may be explored if the algorithm is run more times, with a new choice of the region to explore at each time.

Future developments of this work include the formulation of a dedicated heuristic to solve the problem, so that it can be run in shorter times and possibly in networks of larger dimension. The aim is expanding the use of this model to other types of networks.

Acknowledgements

The authors acknowledge the support of José Salgado, Vitor Mirones and their team from Altice Labs (formerly PT Inovação). The authors would also like to thank the anonymous reviewers for their insightful comments, which helped to improve this paper.

Funding

This work was financially supported by programme COMPETE of the EC Community Support Framework III [QREN 23301 PANORAMA II], co-financed by European Union's FEDER through "Programa Operacional Factores de Competitividade" (POFC) of QREN [FCOMP-01-0202-FEDER-023301], and national funds – Portuguese Foundation for Science and Technology [UID/MULTI/00308/2013].

References

Alouneh, S., Agarwal, A., and En-Nouaary, A., 2009. A novel path protection scheme for MPLS networks using multi-path routing. *Computer Networks*, 53 (9), 1530–1545.

Alvarez, S., 2006. *QoS for IP/MPLS Networks.* Cisco Press.

Banks, J. and Carson, J.S., 1984. *Discrete-Event System Simulation.* International Series in Industrial and Systems Engineering. Prentice-Hall.

Clímaco, J., Craveirinha, J., and Girão-Silva, R., 2016. Multicriteria analysis in telecommunication network planning and design: A survey. *In*: S. Greco, M. Ehrgott and J.R. Figueira, eds. *Multiple Criteria Decision Analysis – State of the Art Surveys*, Vol. 233 of *International Series in Operations Research & Management Science.* Springer, chap. 26, 1167–1233.

Clímaco, J.C.N., Craveirinha, J.M.F., and Pascoal, M.M.B., 2007. Multicriteria routing models in telecommunication networks – Overview and a case study. *In*: Y. Shi, D.L. Olson and A. Stam, eds. *Advances in Multiple Criteria Decision Making and Human Systems Management: Knowledge and Wisdom.* IOS Press, 17–46.

Cohon, J.L., 1978. *Multiobjective Programming and Planning.* Mathematics in Science and Engineering. Academic Press.

Craveirinha, J., Clímaco, J., Martins, L., da Silva, C.G., and Ferreira, N., 2013. A bi-criteria minimum spanning tree routing model for MPLS/Overlay networks. *Telecommunication Systems*, 52 (1), 203–215.

Craveirinha, J., Girão-Silva, R., and Clímaco, J., 2008. A meta-model for multiobjective routing in MPLS networks. *Central European Journal of Operations Research*, 16 (1), 79–105.

Craveirinha, J.M.F., Clímaco, J.C.N., Pascoal, M.M.B., and Martins, L.M.R.A., 2007. Traffic splitting in MPLS networks – A hierarchical multicriteria approach. *Journal of Telecommunications and Information Technology*, (4), 3–10.

Doar, M. and Leslie, I.M., 1993. How bad is naive multicast routing? *In*: *Proceedings of INFOCOM*, Vol. 1, San Francisco (CA), USA, 82–89.

Ehrgott, M. and Gandibleux, X., 2000. A survey and annotated bibliography of multiobjective combinatorial optimization. *OR Spektrum*, 22 (4), 425–460.

Erbas, S.C. and Erbas, C., 2003. A multiobjective off-line routing model for MPLS networks. *In*: J. Charzinski, R. Lehnert and P. Tran-Gia, eds. *Proceedings of the 18th International Teletraffic Congress (ITC-18).* Berlin, Germany: Elsevier, Amsterdam, 471–480.

Fortz, B. and Thorup, M., 2000. Internet traffic engineering by optimizing OSPF weights. *In*: M. Sidi, I. Katzela and Y. Shavitt, eds. *Proceedings of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2000)*, Vol. 2, Mar. 26-30 Tel Aviv, Israel: IEEE Computer and Communications Societies, 519–528.

Fortz, B. and Thorup, M., 2002. Optimizing OSPF/IS-IS weights in a changing world. *IEEE Journal on Selected Areas in Communications*, 20 (4), 756–767.

Girão-Silva, R., Craveirinha, J., Clímaco, J., and Captivo, M.E., 2015. Multiobjective routing in multiservice MPLS networks with traffic splitting – A network flow approach. *Journal of Systems Science and Systems Engineering*, 24 (4), 389–432.

Gomes, T., Martins, L., and Craveirinha, J., 2001. An algorithm for calculating k shortest paths with a maximum number of arcs. *Investigação Operacional*, 21, 235–244.

GT-ITM, 2000. Modeling topology of large internetworks. GT-ITM. Accessed August 19 2016. http://www.cc.gatech.edu/projects/gtitm/

Huang, C., Sharma, V., Owens, K., and Makam, S., 2002. Building reliable MPLS networks using a path protection mechanism. *IEEE Communications Magazine*, 40 (3), 156–162.

IBM, 2013. ILOG CPLEX Optimization Studio V12.6. IBM.

IBM, 2013. ILOG CPLEX Optimization Studio V12.6.0 documentation. IBM. Accessed August 19 2016. http://pic.dhe.ibm.com/infocenter/cosinfoc/v12r6/index.jsp

Jorge, L. and Gomes, T., 2006. Survey of recovery schemes in MPLS networks. *In*: W. Zamojski, J. Mazurkiewicz, J. Sugier and T. Walkowiak, eds. *Proceedings of DepCoS – International Conference on Dependability of Computer Systems (RELCOMEX 2006)*, May 25-27, Szklarska Poreba, Poland, 110–118.

Martins, E.Q.V., Pascoal, M.M.B., and Santos, J.L.E., 1999. Deviation algorithms for ranking shortest paths. *International Journal of Foundations of Computer Science*, 10 (3), 247–263.

Mavrotas, G., 2009. Effective implementation of the ϵ-constraint method in Multi-Objective Mathematical Programming problems. *Applied Mathematics and Computation*, 213 (2), 455–465.

Mitra, D. and Ramakrishnan, K.G., 2001. Techniques for traffic engineering of multiservice, multipriority networks. *Bell Labs Technical Journal*, 6 (1), 139–151.

Mobasheri, B. and Moghadam, M.H.Y., 2014. A path protection method using congestion control in IP/MPLS networks as an underlying network in Smart Grids. *The SIJ Transactions on Computer Science Engineering & its Applications (CSEA)*, 2 (5), 182–189.

Naraghi-Pour, M. and Desai, V., 2008. Loop-free traffic engineering with path protection in MPLS VPNs. *Computer Networks*, 52 (12), 2360–2372.

Pan, P., Swallow, G., and Atlas, A., 2005. Fast Reroute extensions to RSVP-TE for LSP tunnels. *RFC 4090*, Network Working Group.

Sharma, V. and Hellstrand, F., 2003. Framework for Multi-Protocol Label Switching (MPLS)-based recovery. *RFC 3469*, Network Working Group.

Srivastava, S., Agrawal, G., Pióro, M., and Medhi, D., 2005. Determining link weight system under various objectives for OSPF networks using a Lagrangian relaxation-based approach. *IEEE Transactions on Network and Service Management*, 2 (1), 9–18.

Steuer, R.E., 1986. *Multiple Criteria Optimization: Theory, Computation and Application*. Probability and Mathematical Statistics. John Wiley & Sons.

Wang, D. and Ergun, F., 2005. Path protection with pre-identification for MPLS networks. *In*: *Proceedings of the 2nd International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks (QSHINE'05)*, Aug., Lake Vista (FL), USA.

Wierzbicki, A.P. and Burakowski, W., 2011. A conceptual framework for multiple-criteria routing in QoS IP networks. *International Transactions in Operational Research*, 18 (3), 377–399.