

Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Ciência da Computação

Planejamento de Topologia Virtual
com Combinação de Tráfegos em
Redes Óticas Multiplexadas por
Divisão de Comprimento de Onda

Zenilton Kleber Gonçalves do Patrocínio Júnior

Belo Horizonte

2005

Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Ciência da Computação

Planejamento de Topologia Virtual
com Combinação de Tráfegos em
Redes Óticas Multiplexadas por
Divisão de Comprimento de Onda

Tese apresentada ao Programa de Pós-Graduação
em Ciência da Computação da Universidade Federal
de Minas Gerais como requisito parcial para obten-
ção do título de Doutor em Ciência da Computação.

Orientador : Prof. Geraldo Robson Mateus

Zenilton Kleber Gonçalves do Patrocínio Júnior

Belo Horizonte
Dezembro 2005

A meus pais, Zenilton e Edna, pelo amor e por tudo que me ensinam.

AGRADECIMENTOS

Ao Prof. Geraldo Robson Mateus, meu orientador, pela inspiração, suporte e sugestões ao longo dos últimos 15 anos, desde de minha graduação.

Aos membros da banca examinadora, Prof. Anilton Salles Garcia, Prof. Antônio Alfredo Ferreira Loureiro, Prof. Carlos Eduardo Ferreira, Prof. Henrique Pacca Loureiro Luna e Prof. Maurício Guilherme de Carvalho Resende pelas críticas e sugestões feitas a este trabalho.

A minha família pelo suporte constante e pela compreensão sem os quais esse trabalho não teria sido realizado.

Ao Álex, meu irmão querido, pela ajuda fundamental para a minha qualificação.

Ao Paulo, amigo querido e sempre presente, que me apoiou desde do princípio da realização deste trabalho.

Aos bons e “velhos” amigos do Centro Universitário Newton de Paiva, Adriana, Leila e Osvaldo, que tanto me ajudaram ao longo dos anos que trabalhamos juntos durante na fase inicial desde trabalho.

A todos os recentes amigos da Pontifícia Universidade Católica de Minas Gerais e, em especial, aos “companheiros” Silvio e Fabio do Colegiado de Sistemas de Informação, pela paciência e ajuda durante o tempo em que se trabalho estava sendo concluído.

A amiga Lucila cujo convite de trabalho me mostrou novas perspectivas, após 12 anos de experiência docente, e me deu forças e estímulo para finalmente concluir este trabalho.

Aos meus alunos pela paciência nesse último semestre letivo e, especialmente, aos “desorientados” Gustavo e Pedro, pela realização de testes sobre a viabilidade das propostas apresentadas neste trabalho.

As diversas pessoas que direta ou indiretamente contribuíram para execução desse trabalho.

Resumo

Neste trabalho, apresenta-se um estudo profundo sobre o *Traffic Grooming* em redes óticas WDM independentemente da topologia da rede física subjacente.

Uma “nova” formulação natural para o problema, obtida a partir de uma representação estendida para a topologia de rede, é proposta e avaliada. Utiliza-se também de uma representação em camadas para a topologia de rede para se obter uma formulação simplificada que serve de base para o desenvolvimento de vários métodos de resolução do problema.

Além da formalização de diversos limites inferiores baseados no uso da relaxação lagrangeana e da realização de um estudo sobre a estrutura facial do poliedro associado ao conjunto de soluções do problema, diversos métodos de resolução baseados nas abordagens lagrangeana e poliédrica foram implementados e avaliados. Os resultados dos experimentos computacionais apontam para superioridade das abordagens lagrangeanas e, em especial, da heurística lagrangeana proposta para resolução do problema.

Além disso, realizou-se uma investigação preliminar sobre a adequação dos métodos desenvolvidos na resolução de uma versão do problema em que se considere a reconfiguração da rede ao longo de um horizonte de tempo limitado e de outra, em que apenas alguns dos elementos da rede são capazes de realizar *grooming*.

Abstract

In this work, the *Traffic Grooming* problem (**TGP**) in WDM optical networks is explored regardless of underlying physical topology.

A new integer linear program (ILP) is presented and tested. A layered graph representation of the problem is also presented. It is used to reformulate the problem and to obtain a “simplified” ILP. Two distinct approaches – Lagrangian-based and polyhedral approaches – are used in order to solve the problem.

Lagrangian relaxation is used to generate lower bounds for TGP and a study is conducted in order to obtain valid inequalities for the proposed ILP. Several methods based on the two approaches – Lagrangian-based and polyhedral approaches – are implemented and tested. Test results suggest that Lagrangian-based approaches (specially, a Lagrangian-based heuristic) seem to perform better than polyhedral ones.

Moreover, two other distinct versions of **TGP** are discussed. The first one is a new version of **TGP** for a “dynamic-grooming” scenario and the second a version of **TGP** in WDM optical network in which only some nodes have traffic-grooming capability (“sparse-grooming” scenario). A preliminary investigation is conducted and results are presented.

Sumário

Resumo	ix
Abstract	xi
Lista de Abreviaturas	xvii
Lista de Símbolos	xix
Lista de Figuras	xxxix
Lista de Tabelas	xxxiii
1 Introdução	1
1.1 Motivação	2
1.1.1 Aspectos Estruturais	3
1.1.2 Aspectos Conjunturais	5
1.1.3 Um Exemplo Prático	8
1.1.4 Sobre um Enfoque de Otimização	10
1.2 Objetivos	11
1.2.1 Tema	12
1.2.2 Delimitação do Tema	12
1.2.3 Objetivo Geral	12
1.2.4 Objetivos Específicos	12
1.3 Principais Contribuições	13
1.4 Organização do Texto	14
2 Revisão da Literatura	15
2.1 Sobre Projeto de Redes Óticas	15
2.2 Sobre o <i>RWA</i>	18
2.3 Sobre o TGP	20
2.3.1 Técnicas de Multiplexação	20
2.3.2 TGP nas Redes em Anel	22
2.3.3 TGP nas Redes Irregulares	26
2.4 Formulações Existentes para o TGP	30

3	Modelos Matemáticos	39
3.1	<i>Traffic Grooming Problem</i>	41
3.1.1	Representação Estendida	41
3.1.2	Formulação Matemática do TGP	43
3.1.3	Representação em Camadas	46
3.1.4	Formulação Matemática Condensada do TGP	48
3.2	<i>Traffic Grooming</i> Esparso	51
3.2.1	Limitando o Número de GOXCs	51
3.2.2	Alterando a Representação em Camadas	53
3.3	Limitando o Número de <i>Transceivers</i>	54
3.4	<i>Traffic Grooming</i> e Reconfiguração de Rede	55
3.4.1	Formulação Matemática do TGRP	56
3.4.2	Formulação Condensada do TGRP	58
3.4.3	Limitando o Número de GOXCs	60
3.4.4	Limitando o Número de <i>Transceivers</i>	60
3.5	Considerações Finais	60
4	Abordagens Propostas	63
4.1	Algoritmo de <i>Branch-and-Bound</i>	64
4.1.1	Limites Inferiores	65
4.1.2	Limites Superiores	66
4.1.3	Problema Lagrangeano Dual	67
4.1.4	Heurística Lagrangeana	68
4.1.5	Algoritmo de Busca	68
4.2	Algoritmo de <i>Branch-and-Cut</i>	69
4.2.1	<i>Planos-de-Corte</i>	71
4.2.2	Método de <i>Planos-de-Corte</i>	72
4.2.3	Algoritmo de <i>Cut-and-Branch</i>	73
4.2.4	Algoritmo de Busca	73
4.3	Considerações Finais	73
5	Abordagem Lagrangeana do TGP	75
5.1	Limites Inferiores para TGP	75
5.1.1	Relaxação #1	76
5.1.2	Relaxação #2	77
5.1.3	Relaxação #3	81
5.1.4	Relaxação #4	82
5.2	Limite Superior para TGP	82
5.3	Método de Subgradiente para TGP	84
5.3.1	Relaxação #1	85
5.3.2	Relaxação #2	86
5.3.3	Relaxação #3	86
5.3.4	Relaxação #4	87
5.3.5	Direção de Subida	87
5.3.6	Tamanho do Passo	88

5.3.7	Algoritmo de Subgradiente	89
5.4	Algoritmo de <i>Branch-and-Bound</i> para TGP	90
5.4.1	Nó Inicial	91
5.4.2	Nós Intermediários	91
5.4.3	Nó Terminal	93
5.5	Heurística Lagrangeana para TGP	93
5.6	Experimentos Computacionais	94
5.6.1	Resultados da Heurística Lagrangeana	94
5.6.2	Resultados do Algoritmo de <i>Branch-and-Bound</i>	107
5.7	Comparação com Outras Abordagens	113
5.7.1	Avaliação Preliminar	113
5.7.2	Avaliação Qualitativa dos Resultados	119
5.8	Considerações Finais	127
6	Abordagem Poliédrica do TGP	131
6.1	Sobre o Poliedro de Soluções – \mathcal{P}_S	131
6.1.1	Dimensão do Poliedro \mathcal{P}_S	135
6.1.2	Sobre o Poliedro \mathcal{P}_X	135
6.1.3	Sobre o Poliedro \mathcal{P}_Y	137
6.1.4	Sobre o Poliedro \mathcal{P}_Z	140
6.1.5	Considerações sobre o Estudo dos Poliedros	141
6.1.6	Inequações Válidas para \mathcal{P}_S	142
6.2	Avaliação Preliminar dos Cortes	144
6.3	Método de <i>Planos-de-Corte</i>	147
6.4	Algoritmo de <i>Branch-and-Cut</i>	148
6.5	Algoritmo de <i>Cut-and-Branch</i>	149
6.6	Experimentos Computacionais	149
6.6.1	Resultados do Método de <i>Planos-de-Corte</i>	149
6.6.2	Resultados do Algoritmo de <i>Branch-and-Cut</i>	154
6.6.3	Resultados do Algoritmo de <i>Cut-and-Branch</i>	158
6.7	Considerações Finais	159
7	Comparação entre Abordagens do TGP	163
7.1	Descrição dos Experimentos	163
7.2	Avaliação dos Resultados	164
7.3	Considerações Finais	168
8	Variações do TGP	169
8.1	Introdução	169
8.2	Abordagem Lagrangeana do TGRP	169
8.2.1	Limite Inferior para TGRP	169
8.2.2	Limite Superior para TGRP	173
8.2.3	Método de Subgradiente para TGRP	173
8.2.4	Heurística Lagrangeana para TGRP	174

8.2.5	Experimentos Computacionais	174
8.3	Abordagem Lagrangeana do STGP	175
8.3.1	Experimentos Computacionais	176
8.4	Considerações Finais	179
9	Conclusões e Trabalhos Futuros	181
9.1	Conclusões	181
9.2	Trabalhos Futuros	182
	Referências Bibliográficas	185
A	Análise Preliminar	199
A.1	Introdução	199
A.2	Sobre as Soluções	199
A.3	Sobre a Estrutura do Grafo	201
A.4	Consideração Finais	206

Lista de Abreviaturas

ADM	<i>Add-Drop Multiplexer</i>
ATM	<i>Asynchronous Transfer Mode</i>
BLSR	<i>Bidirectional Line-Switched Ring</i>
CAC	Controle de Admissão de Conexão
DXC	<i>Digital Crossconnect</i>
FDM	<i>Frequency Division Multiplexing</i>
GB	Gigabyte
GbE	<i>Gigabit Ethernet</i>
Gbps	Gigabits por segundo
GMPLS	<i>Generalized Multi-Protocol Label Switching</i>
GOXC	<i>Grooming Optical Crossconnect</i>
HDTV	<i>High-Definition Television</i>
IETF	<i>Internet Engineering Task Force</i>
IP	<i>Internet Protocol</i>
LAN	<i>Local Area Network</i>
LTE	<i>Line Terminating Equipment</i>
MAN	<i>Metropolitan Area Network</i>
MB	Megabyte
Mbps	Megabits por segundo
MHz	MegaHertz
MPLS	<i>Multi-Protocol Label Switching</i>
OADM	<i>Optical Add-Drop Multiplexer</i>
OC	<i>Optical Carrier</i>
OEO	Conversão Ótico-Eletro-Ótica
OSPF	<i>Open Shortest Path First</i>

OXC	<i>Optical Crossconnect</i>
Pbps	Petabits por segundo
PLI	Programa Linear Inteiro
POCL	Problema de Otimização Combinatória Linear
PPWDM	<i>Point-to-Point WDM</i>
PSTN	<i>Public Switched Telephone Network</i>
QoS	<i>Quality of Service</i>
RAM	<i>Random Access Memory</i>
RISC	<i>Reduced Instruction Set Computer</i>
RWA	<i>Routing and Wavelength Assignment</i>
SDH	<i>Synchronous Digital Hierarchy</i>
SONET	<i>Synchronous Optical Network</i>
STGP	<i>Sparse Traffic Grooming Problem</i>
STM	<i>Synchronous Transport Module</i>
STS	<i>Synchronous Transport Signal</i>
Tbps	Terabits por segundo
TDM	<i>Time Division Multiplexing</i>
TGP	<i>Traffic Grooming Problem</i>
TGP-MH	<i>Traffic Grooming Problem (Multi-Hop)</i>
TGP-SH	<i>Traffic Grooming Problem (Single-Hop)</i>
TGRP	<i>Traffic Grooming and Reconfiguration Problem</i>
WADM	<i>Wavelength Add-Drop Multiplexer</i>
WAN	<i>Wide Area Network</i>
WDM	<i>Wavelength Division Multiplexing</i>
WGXC	<i>Wavelength-Grooming Crossconnect</i>
WR	<i>Wavelength Routing</i>
WT	<i>Wavelength Translating</i>

Lista de Símbolos

A	Conjunto de arcos da representação estendida $G = (N, A)$, página 43
a	Número de arcos da representação estendida, $a = 2e + \sum_{i \in V} a_i^o$, página 42
\hat{A}_0	Conjunto de arcos utilizados inicialmente durante o cálculo de limites superiores de TGP , página 83
\hat{A}_0^t	Conjunto de arcos utilizados inicialmente durante o cálculo de limites superiores de TGRP , página 173
A^e	Conjunto de arcos entre nós de <i>add-drop</i> e os nós óticos da representação estendida $G = (N, A)$, página 43
A^i	Conjunto de arcos do subgrafo $G^i = (N^i, A^i)$ da representação estendida $G = (N, A)$, página 188
\bar{A}^i	Conjunto de arcos do subgrafo $\bar{G}^i = (\bar{N}^i, \bar{A}^i)$ da representação estendida $G = (N, A)$, página 189
a_i^o	Número de arcos adicionados à representação estendida do nó $i \in V$, $a_i^o = g_i^2 + g_i$, página 42
A_ℓ	Conjunto de arcos da representação em camadas $G_\ell = (N_\ell, A_\ell)$, página 48
$A_\ell^+(i)$	Conjunto de arcos que saem do nó $i \in N_\ell$, veja equação (3.8), página 49
$\tilde{A}_\ell^+(i)$	Conjunto de arcos que saem do nó $i \in (\tilde{N}_\ell \cup \{\eta\})$, veja equação (5.8), página 79
$A_\ell^-(i)$	Conjunto de arcos que chegam no nó $i \in N_\ell$, veja equação (3.9), página 49
$\tilde{A}_\ell^-(i)$	Conjunto de arcos que chegam no nó $i \in (\tilde{N}_\ell \cup \{\eta\})$, veja equação (5.9), página 79
$A_\ell^<$	Conjunto de arcos $\ell \in A_\ell$ para os quais o custo modificado \tilde{c}_ℓ é negativo, página 82

\bar{A}_ℓ	Conjunto de arcos selecionados durante o cálculo de limites inferiores de TGP , página 83
\bar{A}_ℓ^t	Conjunto de arcos selecionados durante o cálculo de limites inferiores de TGRP , página 173
\hat{A}_ℓ	Conjunto de arcos selecionados durante o cálculo de limites superiores de TGP , página 84
A_ℓ^p	Conjunto de arcos $\ell \in A_\ell$ para os quais o custo modificado \tilde{c}_ℓ^p associado ao produto $t_p \in P$ é negativo, página 81
\tilde{A}_ℓ	Conjunto de arcos da representação em camadas modificada $\tilde{G}_\ell = (\tilde{N}_\ell, \tilde{A}_\ell)$, página 78
\tilde{A}_ℓ^η	Conjunto de arcos artificiais acrescentados a representação em camadas modificada $\tilde{G}_\ell = (\tilde{N}_\ell, \tilde{A}_\ell)$, página 79
A°	Conjunto de arcos entre nós óticos da representação estendida $G = (N, A)$, página 43
\bar{A}°	Conjunto de arcos entre nós óticos de diferentes elementos da rede, página 45
A_S	Conjunto de arcos do grafo subjacente $G_S = (N_S, A_S)$, página 187
A_S^i	Conjunto de arcos do grafo subjacente ao subgrafo $G^i = (N^i, A^i)$, página 188
\bar{A}_S^i	Conjunto de arcos do grafo subjacente ao subgrafo $\bar{G}^i = (\bar{N}^i, \bar{A}^i)$, página 189
β_i	Número mínimo de arcos necessários para se transportar todo o tráfego originado em um nó de <i>add-drop</i> $i \in N_\ell^e$, página 78
β_i^t	Número mínimo de arcos necessários para se transportar todo o tráfego originado em um nó de <i>add-drop</i> $i \in N_\ell^e$ no período $t \in T$, página 171
$\bar{\beta}_i^t$	Aumento no número mínimo de arcos necessários para se transportar todo o tráfego originado em um nó de <i>add-drop</i> $i \in N_\ell^e$ no período $t \in T$ em relação ao período anterior, isto é, $\bar{\beta}_i^t = \beta_i^t - \beta_i^{t-1}$, página 172
β_T	Somatório das quantidades mínimas de arcos necessárias para se transportar todo o tráfego originado nos nós de <i>add</i> , $\beta_T = \sum_{i \in \tilde{N}_\ell^a} \beta_i$, página 78
b_i^p	Balanço de fluxo do produto t_p no nó $i \in N_\ell$, veja equação (3.11), página 49

\bar{b}_i^p	Balanço de fluxo do produto t_p no nó $i \in N_\ell$, veja equação (6.2), página 132
b_i^{pt}	Balanço de fluxo do produto t_p no nó $i \in N_\ell$ no período $t \in T$, veja equação (3.20), página 59
b_i^w	Balanço de fluxo do nó $i \in (\tilde{N}_\ell \cup \{\eta\})$ no subproblema L_{2W}^{TGP} , veja equação (5.11), página 79
C	Conjunto de <i>containers virtuais</i> (ou comprimentos de onda), tal que $C = \{1, 2, \dots, W\}$, página 43
c_1	Parcela variável do custo de estabelecimento de um enlace físico, veja equação (3.6), página 45
c_2	Parcela fixa do custo de estabelecimento de um enlace físico, veja equação (3.6), página 45
\mathcal{C}_{cam}	Conjunto de caminhos associados a uma solução viável do TGP , página 186
$\mathcal{C}_{\text{cam}}^p$	Conjunto de caminhos associados a uma solução viável do TGP para um dado produto $t_p \in P$, página 186
$\mathcal{C}_{\text{ciclo}}$	Conjunto de ciclos associados a uma solução viável do TGP , página 186
$\mathcal{C}_{\text{ciclo}}^p$	Conjunto de ciclos associados a uma solução viável do TGP para um dado produto $t_p \in P$, página 186
c_{ijk}	Custo da alocação do <i>container virtual</i> (ou comprimento de onda) $k \in C$ no arco $(i, j) \in A$, página 43
c_{ijk}^t	Custo de utilização do <i>container virtual</i> (ou comprimento de onda) $k \in C$ no arco $(i, j) \in A$ no período $t \in T$, página 56
c_ℓ	Custo da utilização do arco $\ell \in A_\ell$, página 48
\tilde{c}_ℓ^p	Custo modificado do fluxo do produto $t_p \in P$ através do arco $\ell \in A_\ell$ usado no cálculo de L_1^{TGP} , página 76
c_ℓ^t	Custo de utilização do arco $\ell \in A_\ell$ no período $t \in T$, página 58
c_{LTE}	Custo associado a instalação de um LTE, página 45
\tilde{c}_ℓ	Custo modificado do arco $\ell \in A_\ell$ usado no cálculo de L_1^{TGP} , página 76
\mathcal{C}^p	Conjunto de arcos pertencentes ao caminho mínimo entre s_p e d_p calculado utilizando-se λ_ℓ como custo de cada arco, página 80
\mathcal{C}^{pt}	Conjunto de arcos pertencentes ao caminho mínimo entre s_p e d_p calculado utilizando-se λ_ℓ^t como custo de cada arco, página 170

\tilde{c}	Custo de um arco de <i>add-drop</i> , veja equação (3.6), página 45
Δ	Capacidade de um <i>container virtual</i> (ou comprimento de onda), veja equação (3.1), página 43
$\delta_{G_\ell}(N_1, N_2)$	Conjunto de todos os arcos com origem em um nó de N_1 destinados a um nó de N_2 , página 139
$\delta_{G_\ell}^+(N_1)$	Conjunto de todos os arcos cujos nós de origem pertençam a N_1 e os nós de destino não pertençam, isto é, $\delta_{G_\ell}^+(N_1) = \delta_{G_\ell}(N_1, N_\ell \setminus N_1)$, página 143
$\delta_{G_\ell}(N_1, \dots, N_k)$	Conjunto de todos os arcos com origem em um nó de N_i destinados a um nó de N_j , em que $i \neq j$, página 139
δ_p	Contribuição (ou melhor, a ocupação) correspondente a uma unidade de fluxo de um produto transportado em um <i>container virtual</i> (ou comprimento de onda), veja equação (3.2), página 44
$\delta_P(N_1, N_2)$	Conjunto de produtos cujos nós de origem pertencem a N_1 e os nós de destino pertencem a N_2 , página 142
$\delta_P(N_1, \dots, N_k)$	Conjunto de produtos cujos nós de origem e de destino pertencem a subconjuntos distintos da partição N_1, N_2, \dots, N_k , página 139
d^k	Direção de subida gerada a partir de ξ^k na k -ésima iteração, página 87
d_1^k	Direção de subida gerada a partir de ξ_1^{TGP} na k -ésima iteração, veja equação (5.27), página 85
\bar{d}_1^k	Direção de subida gerada a partir de ξ_1^{TGRP} na k -ésima iteração, veja equação (8.13), página 174
d_2^k	Direção de subida gerada a partir de ξ_2^{TGP} na k -ésima iteração, veja equação (5.29), página 86
d_3^k	Direção de subida gerada a partir de ξ_3^{TGP} na k -ésima iteração, veja equação (5.31), página 86
d_4^k	Direção de subida gerada a partir de ξ_4^{TGP} na k -ésima iteração, veja equação (5.33), página 87
d_p	Nó de destino do tráfego t_p , página 41
E	Conjunto de arcos, isto é, conexões entre os elementos da rede, página 41
e	Número de arcos da topologia básica, $e = E $, página 41
ϵ_1	<i>Gap de dualidade</i> mínimo, página 89

ϵ_2	Valor mínimo do parâmetro σ de ajuste do tamanho do passo, página 89
ϵ_3	Valor mínimo da norma do subgradiente, página 89
η	Nó artificial acrescentado à representação em camadas modificada $\tilde{G}_\ell = (\tilde{N}_\ell, \tilde{A}_\ell)$, página 78
ϕ_i	Multiplicadores de Lagrange associados às restrições de continuidade dos <i>lightpaths</i> – restrições (3.10c), página 75
f_{ijk}^p	Variável inteira que representa a quantidade (em canais) do produto $t_p \in P$ que trafega através do arco $(i, j) \in A$ utilizando o <i>container virtual</i> $k \in C$, página 43
f_{ijk}^{pt}	Variável inteira que representa a quantidade (em canais) do produto $t_p \in P$ que trafega através do arco $(i, j) \in A$ utilizando o <i>container virtual</i> $k \in C$ no período $t \in T$, página 56
f_{ik}^p	Variável inteira que representa a quantidade (em canais) do produto $t_p \in P$ inserida/retirada no nó $i \in N^{ad}$ utilizando o <i>container virtual</i> $k \in C$, página 52
f_ℓ^p	Variável inteira que representa a quantidade (em canais) do produto $t_p \in P$ que trafega através do arco $\ell \in A_\ell$, página 48
f_ℓ^{pt}	Variável inteira que representa a quantidade (em canais) do produto $t_p \in P$ que trafega através do arco $\ell \in A_\ell$ no período $t \in T$, página 58
f_p	Volume de tráfego (em número de canais) que deve ser roteado entre s_p e d_p , página 41
\bar{f}_p	Demanda de tráfego entre s_p e d_p na formulação modificada $TGPM$, página 133
f_p^t	Quantidade (em número de canais) do produto $t_p \in P$ que deve ser transportada de s_p a d_p no período $t \in T$, página 56
$G = (N, A)$	Representação estendida, página 43
γ_i	Número mínimo de arcos necessários para se transportar todo o tráfego destinado a um nó de <i>add-drop</i> $i \in N_\ell^e$, página 78
γ_i^t	Número mínimo de arcos necessários para se transportar todo o tráfego destinado a um nó de <i>add-drop</i> $i \in N_\ell^e$ no período $t \in T$, página 171
$\bar{\gamma}_i^t$	Aumento no número mínimo de arcos necessários para se transportar todo o tráfego destinado a um nó de <i>add-drop</i> $i \in N_\ell^e$ no período $t \in T$ em relação ao período anterior, isto é, $\bar{\gamma}_i^t = \gamma_i^t - \gamma_i^{t-1}$, página 172

γ_T	Somatório das quantidades mínimas de arcos necessárias para se transportar todo o tráfego destinado aos nós de <i>drop</i> , $\gamma_T = \sum_{i \in \tilde{N}_\ell^d} \gamma_i$, página 78
$G^i = (N^i, A^i)$	Subgrafo da representação estendida $G = (N, A)$ induzido pelo conjunto de nós óticos associados ao elemento $i \in V$ da rede $H = (V, E)$, página 188
g_i	Grau de nó $i \in V$, página 42
$\bar{G}^i = (\bar{N}^i, \bar{A}^i)$	Subgrafo da representação estendida $G = (N, A)$ induzido pelo nó de <i>add-drop</i> juntamente com os nós óticos associados ao elemento $i \in V$ da rede $H = (V, E)$, página 189
$G_\ell = (N_\ell, A_\ell)$	Representação em camadas, página 48
$\tilde{G}_\ell = (\tilde{N}_\ell, \tilde{A}_\ell)$	Representação em camadas modificada, página 78
$G_S = (N_S, A_S)$	Grafo subjacente a um grafo direcionado $G = (N, A)$, página 187
$G_S^i = (N_S^i, A_S^i)$	Grafo subjacente ao subgrafo $G^i = (N^i, A^i)$, página 188
$\bar{G}_S^i = (\bar{N}_S^i, \bar{A}_S^i)$	Grafo subjacente ao subgrafo $\bar{G}^i = (\bar{N}^i, \bar{A}^i)$, página 189
$H = (V, E)$	Grafo não direcionado representando a topologia básica da rede, página 41
κ	Parâmetro de ajusta da estimativa de solução do problema lagrangeano dual, veja equação (5.40), página 89
L_1^{TGP}	Primeiro limite inferior para TGP obtido através da relaxação lagrangeana, veja equação (5.1), página 76
L_1^{TGRP}	Primeiro limite inferior para TGRP obtido através da relaxação lagrangeana, veja equação (8.1), página 170
L_{1F}^{TGRP}	Subproblema nas variáveis de fluxo associado ao cálculo de L_1^{TGRP} , veja equação (8.1), página 170
$L_{1F_{pt}}^{\text{TGRP}}$	Subproblema em uma única variável de fluxo para um único período associado ao cálculo de L_{1F}^{TGRP} , ver equação (8.2), página 170
L_{1W}^{TGRP}	Subproblema nas variáveis de decisão associado ao cálculo de L_1^{TGRP} , veja equação (8.1), página 170
L_2^{TGP}	Segundo limite inferior para TGP obtido através da relaxação lagrangeana, veja equação (5.6), página 77
L_{2F}^{TGP}	Subproblema nas variáveis de fluxo associado ao cálculo de L_2^{TGP} , veja equação (5.6), página 77

$L_{2F_p}^{\text{TGP}}$	Subproblema em uma única variável de fluxo associado ao cálculo de L_2^{TGP} , veja equação (5.13), página 80
L_{2W}^{TGP}	Subproblema nas variáveis de decisão associado ao cálculo de L_2^{TGP} , veja equação (5.6), página 77
L_3^{TGP}	Terceiro limite inferior para TGP obtido através da relaxação lagrangeana, veja equação (5.17), página 81
L_{3F}^{TGP}	Subproblema nas variáveis de fluxo associado ao cálculo de L_3^{TGP} , veja equação (5.17), página 81
L_{3W}^{TGP}	Subproblema nas variáveis de decisão associado ao cálculo de L_3^{TGP} , veja equação (5.17), página 81
L_4^{TGP}	Quarto limite inferior para TGP obtido através da relaxação lagrangeana, veja equação (5.19), página 82
L_{4F}^{TGP}	Subproblema nas variáveis de fluxo associado ao cálculo de L_4^{TGP} , veja equação (5.19), página 82
L_{4W}^{TGP}	Subproblema nas variáveis de decisão associado ao cálculo de L_4^{TGP} , veja equação (5.19), página 82
λ_ℓ	Multiplicadores de Lagrange associados às restrições que limitam a capacidade dos arcos – restrições (3.10d), página 75
λ_ℓ^t	Multiplicadores de Lagrange associados às restrições que limitam a capacidade dos arcos – restrições (3.19d), página 169
\bar{L}	Estimativa da solução do problema lagrangeano dual, veja equação (5.40), página 89
\hat{L}_{\min}	Menor limite superior encontrado, página 89
l_{ij}	Comprimento “real” do enlace utilizado na ligação entre os elementos de rede i e j , veja equação (3.6), página 45
L^k	Valor do limite inferior calculado na k -ésima iteração, veja equação (5.38), página 88
L_{\max}	Maior limite inferior encontrado, página 89
\mathcal{LP}_S	Poliedro obtido a partir da relaxação linear das variáveis de fluxo de TGP , página 131
\mathcal{LP}_Z	Poliedro associado a relaxação linear do conjunto de soluções Z , página 138
\mathcal{M}	Conjunto de soluções da formulação modificada TGP_M , página 132

μ	Parâmetro de ajuste da <i>regra de Camerini–Fratta–Maffioli</i> , veja equação (5.36), página 88
μ_ℓ	Peso não negativo associado ao arco $\ell \in A_\ell$, página 139
m_p	Número máximo de canais que podem ser transportados em um comprimento de onda, página 41
N	Conjunto de nós da representação estendida $G = (N, A)$, $N = N^e \cup N^o$, página 43
n	Número de nós da representação estendida, $n = v + \sum_{i \in V} n_i^o$, página 42
N^{ad}	Conjunto de nós que realizam apenas <i>add-drop</i> , página 51
N^e	Conjunto de nós de <i>add-drop</i> da representação estendida $G = (N, A)$, página 43
N^g	Conjunto de nós que realizam <i>grooming</i> além de realizarem <i>add-drop</i> , página 51
N^i	Conjunto de nós do subgrafo $G^i = (N^i, A^i)$ da representação estendida $G = (N, A)$, página 188
\bar{N}^i	Conjunto de nós do subgrafo $\bar{G}^i = (\bar{N}^i, \bar{A}^i)$ da representação estendida $G = (N, A)$, página 189
n_i^o	Número de nós óticos acrescentados à representação estendida do nó $i \in V$, $n_i^o = 2g_i$, página 42
N_ℓ	Conjunto de nós da representação em camadas $G_\ell = (N_\ell, A_\ell)$, $N_\ell = N_\ell^e \cup N_\ell^o$, página 48
N_L	Número total de <i>lightpaths</i> , veja equação (3.5), página 45
\tilde{N}_ℓ^a	Conjunto de nós de <i>add</i> da representação em camadas modificada \tilde{G}_ℓ , página 78
\tilde{N}_ℓ^d	Conjunto de nós de <i>drop</i> da representação em camadas modificada \tilde{G}_ℓ , página 78
N_ℓ^e	Conjunto de nós de <i>add-drop</i> da representação em camadas $G_\ell = (N_\ell, A_\ell)$, página 48
N_ℓ^o	Conjunto de nós óticos da representação em camadas $G_\ell = (N_\ell, A_\ell)$, página 48
\tilde{N}_ℓ^o	Conjunto de nós óticos da representação em camadas modificada \tilde{G}_ℓ , página 78
\tilde{N}_ℓ	Conjunto de nós da representação em camadas modificada $\tilde{G}_\ell = (\tilde{N}_\ell, \tilde{A}_\ell)$, tal que $\tilde{N}_\ell = \tilde{N}_\ell^a \cup \tilde{N}_\ell^d \cup \tilde{N}_\ell^o$, página 78

N^o	Conjunto de nós óticos da representação estendida $G = (N, A)$, página 43
n_p	Número total de tráfegos (ou produtos), $n_p = P $, página 41
N_i^R	Número de receptores óticos existentes no elemento da rede associado ao nó de <i>add-drop</i> $i \in N^e$, página 54
N_S	Conjunto de nós do grafo subjacente $G_S = (N_S, A_S)$, página 187
N_S^i	Conjunto de nós do grafo subjacente ao subgrafo $G^i = (N^i, A^i)$, página 188
\bar{N}_S^i	Conjunto de nós do grafo subjacente ao subgrafo $\bar{G}^i = (\bar{N}^i, \bar{A}^i)$, página 189
n_t	Número total de períodos, página 56
N_i^T	Número de transmissores óticos existentes no elemento da rede associado ao nó de <i>add-drop</i> $i \in N^e$, página 54
P	Conjunto de tráfegos (ou produtos), $P = \{t_p \mid p = 1, \dots, n_p\}$, página 41
P_1^D	Problema lagrangeano dual associado ao primeiro limite inferior de TGP – L_1^{TGP} , veja equação (5.22), página 84
\bar{P}_1^D	Problema lagrangeano dual associado ao primeiro limite inferior de TGRP – L_1^{TGRP} , veja equação (8.11), página 173
P_2^D	Problema lagrangeano dual associado ao segundo limite inferior de TGP – L_2^{TGP} , veja equação (5.23), página 84
P_3^D	Problema lagrangeano dual associado ao terceiro limite inferior de TGP – L_3^{TGP} , veja equação (5.24), página 84
P_4^D	Problema lagrangeano dual associado ao quarto limite inferior de TGP – L_4^{TGP} , veja equação (5.25), página 84
π_i^p	Multiplicadores de Lagrange associados às restrições de conservação de fluxo dos nós – restrições (3.10b), página 75
π_p^μ	Tamanho do caminho mínimo em G_ℓ entre s_p e d_p obtido utilizando-se μ_ℓ como comprimento do arco $\ell, \forall \ell \in A_\ell$, página 139
\mathcal{P}_M	Poliedro associado a envoltória convexa do conjunto \mathcal{M} de soluções da formulação modificada TGP_M , página 132
\mathcal{P}_S	Poliedro formado pela envoltória convexa das soluções do TGP , página 131
\mathcal{P}_W	Poliedro associado ao conjunto \mathcal{W} de soluções da formulação estendida TGP_E , página 133

$\mathcal{P}_\mathcal{X}$	Poliedro associado ao conjunto \mathcal{X} de soluções do problema de roteamento de fluxo não-capacitado, página 133
$\mathcal{P}_\mathcal{Y}$	Poliedro associado ao conjunto \mathcal{Y} de frações da demanda são roteadas com sucesso através de um problema de fluxo não-capacitado de vários produtos, página 133
$\mathcal{P}_{\bar{\mathcal{Y}}}$	Poliedro associado ao conjunto $\bar{\mathcal{Y}}$ de soluções do problema de fluxo capacitado de vários produtos, página 137
$\mathcal{P}_\mathcal{Z}$	Poliedro associado ao conjunto \mathcal{Z} de soluções de um problema de instalação de capacidades em uma rede sujeita ao fluxo de vários produtos em que se admite somente a possibilidade de se instalar nos arcos da rede capacidades em quantidades inteiras, página 133
$\mathcal{P}_{\mathcal{Z}_\ell}$	Poliedro associado ao conjunto \mathcal{Z}_ℓ de soluções do problema de instalação de uma capacidade inteira em um arco ℓ de uma dada rede, página 140
Q	Subconjunto dos produtos, isto é, $Q \subseteq P$, página 141
ρ^k	Parâmetro que define o grau de influência da direção de subida anterior na direção atual, veja equação (5.34), página 87
r_i^t	Quantidade de facilidades instaladas nos arcos que chegam em $i \in N_\ell^e$ no período $t \in T$ e que será utilizada em um período futuro $t' > t$, página 171
S	Conjunto de soluções viáveis do TGP , página 185
σ^k	Parâmetro de ajuste do tamanho do passo na k -ésima iteração, veja equação (5.38), página 88
s_i^t	Quantidade de facilidades instaladas nos arcos que saem de $i \in N_\ell^e$ no período $t \in T$ e que será utilizada em um período futuro $t' > t$, página 171
s_p	Nó de origem do tráfego t_p , página 41
T	Conjunto de períodos, $T = \{1, 2, \dots, n_t\}$, página 56
τ	Parâmetro de ajuste de limite superior baseado no melhor limite inferior encontrado, página 89
θ_ℓ	Multiplicadores de Lagrange associados às restrições que limitam a alocação de containers – restrições (3.19e), página 169
t^k	Tamanho do passo na k -ésima iteração, veja equação (5.38), página 88
t_p	Tráfego entre um par de nós, $t_p = (s_p, d_p, f_p, m_p), \forall p = 1, \dots, n_p$, página 41

V	Conjunto de nós, isto é, elementos da rede, página 41
v	Número de nós da topologia básica, $v = V $, página 41
$v_{\bar{c}}$	Número de nós em um caminho \bar{c} na rede original $H = (V, E)$, página 186
v_w	Número de nós de um caminho na rede original que utilizam o roteamento por comprimento de onda, página 186
W	Número de <i>containers virtuais</i> (ou comprimentos de onda) disponíveis, página 43
\mathcal{W}	Conjunto de soluções da formulação estendida TGP_E , página 133
w_{ijk}	Variável binária que indica o uso do <i>container virtual</i> $k \in C$ no arco $(i, j) \in A$ no transporte de algum produto, página 43
w_{ijk}^t	Variável binária que indica o uso do <i>container virtual</i> $k \in C$ no arco $(i, j) \in A$ no transporte de algum produto no período $t \in T$, página 56
w_ℓ	Variável binária que indica o uso do arco $\ell \in A_\ell$ no transporte de algum produto, página 48
w_ℓ^t	Variável binária que indica o uso do arco $\ell \in A_\ell$ no transporte de algum produto no período $t \in T$, página 58
\mathcal{X}	Conjunto de soluções do problema de roteamento de fluxo não-capacitado de vários produtos, página 133
ξ_1^{TGP}	Subgradiente associado ao primeiro limite inferior de $\mathbf{TGP} - L_1^{\text{TGP}}$, página 85
ξ_1^{TGRP}	Subgradiente associado ao primeiro limite inferior de $\mathbf{TGRP} - L_1^{\text{TGRP}}$, página 174
ξ_2^{TGP}	Subgradiente associado ao segundo limite inferior de $\mathbf{TGP} - L_2^{\text{TGP}}$, página 86
ξ_3^{TGP}	Subgradiente associado ao terceiro limite inferior de $\mathbf{TGP} - L_3^{\text{TGP}}$, página 86
ξ_4^{TGP}	Subgradiente associado ao primeiro limite inferior de $\mathbf{TGP} - L_4^{\text{TGP}}$, página 87
ξ^k	Subgradiente calculado na k -ésima iteração, página 87
x_ℓ^p	Variável que representa a ocupação da capacidade do arco ℓ pelo fluxo f_ℓ^p , isto é, $x_\ell^p = \delta_p f_\ell^p, \forall t_p \in P, \forall \ell \in A_\ell$, página 131

\mathcal{Y}	Conjunto de frações da demanda são roteadas com sucesso através de um problema de fluxo não-capacitado de vários produtos, página 133
$\bar{\mathcal{Y}}$	Conjunto de soluções do problema de fluxo capacitado de vários produtos, página 137
\bar{y}_ℓ	Variável que representa capacidade total a ser instalada no arco $\ell \in A_\ell$, página 137
y_ℓ^p	Variável que representa a fração da demanda \bar{f}_p do produto t_p transportada através do arco ℓ , de modo que $x_\ell^p = \bar{f}_p y_\ell^p, \forall t_p \in P, \forall \ell \in A_\ell$, página 133
\mathcal{Z}	Conjunto de soluções de um problema de instalação de capacidades em uma rede sujeita ao fluxo de vários produtos em que se admite somente a possibilidade de se instalar nos arcos da rede capacidades em quantidades inteiras, página 133
\mathcal{Z}_ℓ	Conjunto de soluções do problema de instalação de uma capacidade inteira em um arco ℓ de uma dada rede, página 140
z_ℓ	Variável que representa a capacidade total instalada no arco ℓ , isto é, $z_\ell = \Delta w_\ell, \forall \ell \in A_\ell$, página 133

Lista de Figuras

1.1	Exemplo Prático de <i>Traffic Grooming</i>	9
2.1	Arquiteturas de Elementos de Anel SONET+WDM	22
2.2	Arquitetura <i>Single-Hop</i> \times <i>Multi-Hop</i>	24
2.3	Exemplo de Anéis Interconectados	25
2.4	Arquiteturas dos Elementos de Interconexão	26
2.5	Arquitetura de OXC que realiza <i>Grooming</i>	28
3.1	Subdivisão de um Nó	42
3.2	Exemplo da Representação Estendida	43
3.3	Representações de uma Rede com 2 Nós	47
3.4	Representação em Camadas para STGP	54
5.1	Método de subgradientes	86
5.2	Esquema genérico da Heurística Lagrangeana	93
5.3	Redes com Topologia Irregular	95
5.4	Rede Ótica de Transporte Pan-Européia	95
5.5	Gap Médio das Redes em Anel	97
5.6	Gap Médio para Relaxações #2 e #3 das Redes em Anel	99
5.7	Gap Médio das Redes Irregulares	100
5.8	Gap Médio para Relaxações #2 e #3 das Redes Irregulares	102
5.9	Gap Médio das Variações das Relaxações para Redes em Anel	105
5.10	Gap Médio das Variações das Relaxações para Redes Irregulares	106
5.11	Rede em Anel com 4 Nós	113
5.12	Solução Obtida para TGP com $m_p = 48$	117
5.13	Solução Obtida para TGP com $m_p = 12$	117
5.14	Solução Obtida para TGP com $m_p = 3$	118
5.15	Exemplo de Grafo Auxiliar (adaptado de [166])	119
5.16	Comparação dos Resultados Médios dos Algoritmos de <i>Branch-and-Bound</i> para Redes em Anel	125
5.17	Comparação dos Resultados Médios dos Algoritmos de <i>Branch-and-Bound</i> para Redes Irregulares	126
5.18	Comparação dos Resultados Médios da Heurística Lagrangeana para Redes em Anel	128
5.19	Comparação dos Resultados Médios da Heurística Lagrangeana para Redes Irregulares	129

6.1	Comparação dos Resultados Médios do Método de <i>Planos-de-Corte</i> para Redes em Anel	152
6.2	Comparação dos Resultados Médios do Método de <i>Planos-de-Corte</i> para Redes Irregulares	153
6.3	Comparação dos Resultados Médios do Algoritmo de <i>Branch-and-Cut</i> para Redes em Anel	156
6.4	Comparação dos Resultados Médios do Algoritmo de <i>Branch-and-Cut</i> para Redes Irregulares	157
6.5	Comparação dos Resultados Médios do Algoritmo de <i>Cut-and-Branch</i> para Redes em Anel	160
6.6	Comparação dos Resultados Médios do Algoritmo de <i>Cut-and-Branch</i> para Redes Irregulares	161
7.1	Comparação dos Resultados entre Abordagens Heurísticas	165
7.2	Comparação dos Resultados entre Abordagens Exatas	167
8.1	Teste da Heurística Lagrangeana para STGP	177

Lista de Tabelas

1.1	Matriz de Tráfego do Exemplo	8
5.1	Resultados da Heurística Lagrangeana para Redes em Anel (valores médios)	96
5.2	Resultados da Heurística Lagrangeana para Redes Irregulares (valores médios)	98
5.3	Resultados da Heurística Lagrangeana para a Rede Pan-Européia da Fig. 5.4	101
5.4	Resultados das Variações das Relaxações #2 e #3 do TGP para Redes em Anel (valores médios)	103
5.5	Resultados das Variações das Relaxações #2 e #3 do TGP para Redes Irregulares (valores médios)	104
5.6	Resultados do Algoritmo de <i>Branch-and-Bound</i> BBR2 para Redes em Anel (valores médios)	108
5.7	Resultados do Algoritmo de <i>Branch-and-Bound</i> BBR3 para Redes em Anel (valores médios)	109
5.8	Resultados do Algoritmo de <i>Branch-and-Bound</i> BBR2 para Redes Irregulares (valores médios)	111
5.9	Resultados do Algoritmo de <i>Branch-and-Bound</i> BBR3 para Redes Irregulares (valores médios)	112
5.10	Resultados do Algoritmo de <i>Branch-and-Bound</i> BBR2 para Rede Pan-Européia	114
5.11	Resultados do Algoritmo de <i>Branch-and-Bound</i> BBR3 para Rede Pan-Européia	115
5.12	Tráfego do Anel com 4 Nós	116
5.13	Comparação entre Abordagens	116
5.14	Custos dos Arcos do Grafo Auxiliar para cada Política de <i>Grooming</i> [166]	121
5.15	Custos dos Arcos do Grafo Auxiliar da Política MinWL*	122
5.16	Resultados da Heurística proposta em [166] (valores médios)	123
6.1	Resultado dos <i>Cortes</i> nas Redes em Anel (valores médios)	145
6.2	Resultado dos <i>Cortes</i> nas Redes Irregulares (valores médios)	146
6.3	Resultado dos <i>Cortes</i> na Rede Pan-Européia da Fig. 5.4	147
6.4	Resultado dos <i>Cortes</i> para o Anel da Seção 5.7	147
6.5	Resultados do Método de <i>Planos-de-Corte</i> para Redes em Anel (valores médios)	150

6.6	Resultados do Método de <i>Planos-de-Corte</i> para Redes Irregulares (valores médios)	151
6.7	Resultados do Algoritmo de <i>Branch-and-Cut</i> para Redes em Anel (valores médios)	154
6.8	Resultados do Algoritmo de <i>Branch-and-Cut</i> para Redes Irregulares (valores médios)	155
6.9	Resultados do Algoritmo de <i>Cut-and-Branch</i> para Redes em Anel (valores médios)	158
6.10	Resultados do Algoritmo de <i>Cut-and-Branch</i> para Redes Irregulares (valores médios)	159
8.1	Resultados do TGRP para Redes em Anel (valores médios)	175
8.2	Resultados do TGRP para Redes Irregulares (valores médios)	176
8.3	Resultados do STGP para Redes em Anel (valores médios)	178
8.4	Resultados do STGP para Redes Irregulares (valores médios)	179

Capítulo 1

Introdução

NOS últimos anos, tanto a área de telecomunicações quanto a de redes de computadores têm experimentado um enorme crescimento. Com a popularidade crescente da *Internet* e da *World Wide Web*, e também em virtude da recente privatização e crescente desregulamentação do setor de telecomunicações não só no Brasil, mas por todo o mundo, tal crescimento deve ser observado ainda por um longo período. Nas próximas décadas assistir-se-á a um aumento do número de conexões e do tráfego não só de caráter empresarial mas também domiciliar, de modo a fornecer uma nova gama de serviços, tais como: HDTV (*High-Definition Television*), *video-mail*, áudio digital; além de conexões mais velozes e baratas para acesso a *Internet*. Na medida que houver um aumento do número de usuários das redes de voz/dados e que os mesmos passem a fazer uso com maior frequência de aplicações com alto consumo de banda, haverá uma carência cada vez maior de mecanismos de transporte de alta capacidade e escalabilidade, cujo volume excede em muito ao que já é disponibilizado pelas técnicas adotadas nas atuais redes de alto desempenho, tais como GbE (*Gigabit Ethernet*), ATM (*Asynchronous Transfer Mode*), SONET/SDH (*Synchronous Optical Network/Synchronous Digital Hierarchy*) e GMPLS (*Generalized Multi-Protocol Label Switching*)

Sob esse prisma, a chave para o futuro das redes reside – ao que tudo indica – no campo relativamente recente das redes óticas de transporte. Uma fibra ótica fornece uma grande largura de banda de transmissão, com baixa taxa de perda e um custo relativamente modesto que viabilizará o transporte de voz/dados na escala e capacidade descritos anteriormente, permitindo com isto o atendimento da demanda futura. Dado que a largura em potencial de uma única fibra é enorme, muitos esforços têm sido realizados no sentido de se explorar essa capacidade da melhor forma.

A multiplexação por divisão de comprimento de onda – WDM (*Wavelength Division Multiplexing*) tem se mostrado como uma das mais promissoras abordagens para se explorar o potencial das fibras óticas como meio de transporte. Na WDM, o espectro ótico de transmissão é dividido em várias faixas de comprimento de onda (ou frequência) não sobrepostas, de modo que cada uma delas funciona como um canal separado de comunicação capaz de operar na velocidade de “pico” do processamento eletrônico. Uma vez que, através do uso da WDM, múltiplos canais de

comunicações podem coexistir e operar em uma única fibra, novos desafios (e linhas de investigação) se colocam à nossa frente, tais como definição de novos protocolos e algoritmos de roteamento, planos de controle e gerência de redes, bem como aqueles relacionados ao planejamento e projeto de arquiteturas de rede adequadas a essa nova realidade, principalmente em face dos recentes avanços tecnológicos e, até mesmo, daqueles não disponíveis plenamente mas já previstos para um futuro não muito distante, como a utilização de elementos de rede (roteadores e comutadores) totalmente óticos nas chamadas redes puramente óticas (*All-Optical Networks* ou *Transparent Networks*) ou mesmo na presença de apenas alguns elementos com capacidades limitadas de processamento eletro-eletrônico nas redes híbridas (*Translucent Networks*) [45].

1.1 Motivação

Parafraseando Tanenbaum [145], o significado real da largura de banda *infinita*¹ (apesar do alto custo) ainda não foi totalmente assimilado e o pensamento dominante ainda é determinado pelos limites teóricos de Nyquist e Shannon [140, 145], impostos pelo uso de fios de cobre como meio de transmissão. Contudo, com o desenvolvimento e a utilização de redes óticas – em que luz é utilizada como portadora e as fibras óticas, como meio de transmissão – novos conceitos têm surgido, principalmente, segundo o mesmo autor, em virtude da premissa de que todos os computadores (baseados em mecanismos eletro-eletrônicos) são extremamente lentos e, dessa forma, as redes de uma maneira geral devem evitar qualquer computação (execução de algoritmos via software e/ou via hardware) a todo custo, independentemente de um eventual desperdício de largura de banda.

As redes óticas têm sido amplamente exploradas nas ligações de longa distância (*Wide Area Networks* – *WANs*) e em grandes áreas metropolitanas (*Metropolitan Area Networks* – *MANs*) principalmente na interligação das centrais (centros de fios ou de comutação) das operadoras de telecomunicação. Até bem pouco tempo atrás, a sua utilização em redes locais (*Local Area Networks* – *LANs*) foi limitada principalmente devido ao elevado custo de implantação dos meios físicos de transmissão se comparado ao custo de outras tecnologias disponíveis no mercado. Contudo, a necessidade crescente de uma melhora na qualidade de serviço (*Quality of Service* – *QoS*) prestado aos usuários finais – dada a ampla variedade de perfis de tráfego hoje existentes – leva a crer que, mesmo que não se adote em um futuro próximo o uso das fibras óticas nas redes locais, elas estarão cada vez mais próximas dos usuários finais de modo a viabilizar o transporte do tráfego crescente que será gerado por estes. Deve-se, porém, mencionar problemas tais como a dificuldade de emenda das fibras, além do fato de não serem dobráveis, como fatores que não só contribuem para o aumento de seu custo de instalação, mas que também podem vir a inviabilizar sua implantação em determinados trajetos (principalmente em redes locais).

Os fatores que motivam o presente trabalho podem ser agrupados em dois conjuntos. No primeiro deles, encontram-se os aspectos ligados à modificação da estrutura

¹Vale dizer que a largura de banda de uma fibra não é verdadeiramente *infinita*. Contudo um estudo recente [110] demonstra que ela é da ordem de 100 Tbps (Terabits por segundo) – podendo em termos práticos e em virtude de sua magnitude ser tratada como *infinita*.

de funcionamento das redes óticas e, em particular, à evolução natural da tecnologia utilizada e que, conseqüentemente, acaba sendo responsável pela necessidade de adaptação das atuais abordagens de gerência e alocação de capacidades nas redes óticas. Em um segundo conjunto, estão fatores importantes relacionados à conjuntura atual e, em especial, ligados a aspectos econômicos e à situação do mercado de redes óticas no início deste século.

As duas subseções seguintes abordam tais aspectos estruturais e conjunturais, respectivamente, procurando relacioná-los ao tema deste trabalho. Em seguida um exemplo prático é apresentado para ilustrar informalmente os “ganhos” obtidos no planejamento de redes óticas a partir da utilização dos resultados produzidos neste trabalho. Finalmente, sob a ótica da otimização, discute-se a necessidade da utilização de abordagens sistemáticas e formais que apresentem não só melhor desempenho (e/ou qualidade nas soluções encontradas) mas que permitam mensurar e avaliar os resultados obtidos.

1.1.1 Aspectos Estruturais

As fibras óticas apresentam uma série de vantagens amplamente conhecidas [23, 145]. Primeiramente, deve-se destacar a baixa atenuação que representa em verdade uma economia significativa para as conexões de longa distância dada a redução do número de repetidores/regeneradores de sinal necessários. As fibras apresentam, ainda, a vantagem de não serem afetadas por interferências eletromagnéticas ou mesmo queda no fornecimento de energia (exceto em suas extremidades onde a luz é emitida e/ou captada). Além disso, elas são imunes à ação corrosiva de vários elementos químicos presentes no ar e, conseqüentemente, se adaptam muito bem às mais diversas situações e/ou regiões, incluindo áreas industriais.

Além disso, para os projetistas e implementadores de redes, as fibras apresentam outras duas grandes vantagens: elas são finas e leves – o que contribui para o barateamento da instalação de novas rotas e, até mesmo, da substituição das antigas; por fim, se mostram muito mais seguras contra eventuais invasões ativas ou passivas (escutas ou “grampos”) dos sistemas de comunicação, uma vez que não desperdiçam luz e dificilmente são interceptadas. “A razão para que a fibra seja melhor do que o cobre é inerente às questões físicas subjacentes a esses dois materiais” [145]. Ao se utilizar o cobre, os elétrons dentro do fio afetam e são afetados ao se deslocarem – devido à interação eletromagnética – não só por outros elétrons do próprio fio, mas também por aqueles existentes fora deste. Já na fibra, tais interferências são minimizadas, uma vez que os fótons são totalmente desprovidos de carga.

Ainda segundo [145], outro fator de grande importância no planejamento e projeto de sistemas de comunicação são as economias de escala. Notadamente o custo para instalação e manutenção (ou melhor, os custos decorrentes da instalação propriamente e não do uso de cabos de cobre ou fibras óticas) de um tronco (ligação) de *banda larga* ou de *banda estreita* entre duas centrais de comutação (centro de fios) nas redes telefônicas é praticamente o mesmo [145]. Como conseqüência as operadoras de telecomunicações têm desenvolvido ao longo dos anos – juntamente com a comunidade acadêmica – esquemas elaborados para se multiplexar muitas conversações em um único tronco (ligação) físico, podendo se observar a ocorrência

do mesmo em redes de computadores.

Entre tais esquemas de multiplexação duas categorias básicas merecem destaque: a multiplexação por divisão de frequência – FDM (*Frequency Division Multiplexing*) e a multiplexação por divisão de tempo – TDM (*Time Division Multiplexing*). Na primeira, o espectro de frequência eletromagnético é dividido em canais lógicos em que cada usuário (ou grupo de usuários) tem a posse exclusiva de uma faixa de frequência. Já na segunda, os usuários (ou grupos de usuários) se revezam e cada um deles periodicamente obtém (ou melhor, pode utilizar) toda a largura da banda de comunicação por um determinado período de tempo. No caso das redes óticas e, mais especificamente, de canais de fibra ótica, utiliza-se uma variação da multiplexação por divisão de frequência, conhecida como multiplexação por divisão de comprimento de onda – WDM (*Wavelength Division Multiplexing*). Não há efetivamente nada de novo em tal processo, contudo, diferentemente do que ocorre com a FDM “eletromagnética”, a WDM pode ser realizada por um sistema passivo² (como uma grade de difração) e, portanto, consegue ser altamente veloz e confiável. Os sistemas do início do século XXI já eram capazes de fornecer cerca de uma centena de canais de 10 Gbps e a introdução de melhorias tecnológicas indicam que esta capacidade tende a aumentar já em um futuro bem próximo [41]³.

Em verdade, nos últimos anos uma série de equipamentos para conexão ponto-a-ponto utilizando a tecnologia WDM têm sido lançados no mercado e utilizados na implementação da infraestrutura de diversas redes, principalmente de longa distância (incluindo a própria *Internet*). O grande aumento da largura de banda proporcionado pela WDM tem demonstrado a necessidade da adoção de mecanismos de comutação mais rápidos e eficientes, principalmente entre os elementos centrais das redes, *backbone elements* (aqueles por onde passa a maior parte do tráfego). Mais ainda, a rápida evolução das tecnologias utilizadas nas redes óticas tem possibilitado uma migração sistemática e paulatina dos sistemas de transmissão WDM ponto-a-ponto para uma estrutura em que um *backbone* totalmente ótico é formado por um conjunto de comutadores óticos (*Optical Cross-Connects* – OXCs).

Cada OXC é capaz de comutar um sinal ótico que chega em um dado comprimento de onda através de uma fibra de entrada para o mesmo comprimento de onda em uma fibra ótica de saída através do roteamento por comprimento de onda (*Wavelength Routing* – WR). Um OXC pode, ainda, ser equipado com conversores de modo a permitir que o comprimento de onda de um sinal ótico recebido em uma fibra de entrada seja substituído por outro quando esse sinal é enviado através da fibra de saída (*Wavelength Translating* – WT). O principal mecanismo de transporte nesse tipo de rede é denominado *lightpath* que representa um canal de comunicação estabelecido através da rede de OXCs entre dois elementos situados na “borda” (ou fronteira) da rede de comutadores óticos, podendo passar por várias fibras (enlaces físicos – *physical hops*).

O lançamento comercial de sistemas WDM tem tornado aparente o fato de que

²Sem a utilização de processamento eletro-eletrônico.

³A despeito dos recentes resultados sobre o limite da largura de banda de transmissão de uma fibra [110], com a tecnologia já disponível no início do século XXI, uma boa fonte de laser era capaz de emitir 10^{16} fótons por segundo enquanto que um bom detector – capaz de captar um bit a cada 10 fótons – possibilitaria, pelo menos em tese, uma largura da banda de transmissão da ordem de 1 Pbps (10^{15} bits por segundo) em uma única fibra ótica [41].

o custo dos elementos de rede, em especial dos equipamentos de terminação de linha (*Line Terminating Equipments – LTEs*), dominam os demais custos de “construção” de uma rede ótica [45]. Dessa forma, o número de *LTEs* representa uma métrica mais significativa a se otimizar que, por exemplo, o número de comprimentos de onda disponíveis e/ou utilizados, apesar de no passado (não muito remoto) esse último ter sido geralmente empregado como métrica padrão e a maioria dos trabalhos de pesquisa realizados, até então, ter buscado a sua redução.

Além disso, utilizando a tecnologia do início deste século, cada comprimento de onda pode operar a uma taxa da ordem de 2,5 a 10 Gbps (Gigabits por segundo), ao passo que canais (comprimentos de onda) operando a 40 Gbps estarão disponíveis comercialmente em um futuro próximo. Contudo, a necessidade de banda associada a um tráfego individualmente (uma conexão de dados ou voz) é, via de regra, muito pequena se comparada a tais valores. Daí surge o conceito de *Traffic Grooming*, que se refere ao uso de técnicas de multiplexação para combinar tráfegos de baixa velocidade (que necessitam de pouca largura de banda), de modo a utilizar os canais (comprimentos de onda) disponíveis da melhor forma, ou, ainda, procurando-se atender a um determinado objetivo durante o projeto da rede como minimizar seus custos de implantação e/ou configuração.

A partir da segunda metade da década de 90, o problema de *Traffic Grooming* (*Traffic Grooming Problem – TGP*) tem recebido considerável atenção na literatura especialmente enfocando redes com topologia em anel, uma vez que o emprego de anéis óticos SONET/SDH tem se tornado muito comum desde então. Alguns desses estudos procuram minimizar o custo diretamente a partir da minimização da quantidade de multiplexadores de inserção-retirada (*Add-Drop Multiplexers – ADMs*) utilizados na rede SONET/SDH; enquanto outros o fazem indiretamente através da redução dos custos associados a comutação (roteamento) eletro-eletrônica.

Mais recente, o **TGP** em redes com topologia física irregular vem despertando o interesse da comunidade científica. Isto se deve não só à melhor aderência deste tipo de topologia à realidade, mas principalmente ao fato de tal topologia apresentar uma melhor escalabilidade quando comparada a uma rede em anel. Além disso, como muitas das redes óticas atuais apresentam topologia física em anel, é razoável esperar que em um futuro não muito distante tais anéis sejam conectados uns aos outros de modo a formar uma rede de anéis interconectados (em verdade, já é possível se encontrar tais situações, por exemplo, em grandes regiões metropolitanas).

1.1.2 Aspectos Conjunturais

Embora o mercado global de telecomunicações tenha reduzido de tamanho no início deste século⁴, as redes óticas estão se tornando o “centro nervoso” de nossas redes de telecomunicações. Para se entender de que forma isto vem acontecendo bem como suas causas, é necessário que se faça uma análise das causas do *crash* do mercado de telecomunicações [151]. A combinação perfeita de dois fenômenos foi responsável pelo mencionado *crash*: a desregulamentação generalizada do mercado

⁴Fenômeno denominado por alguns analistas do mercado como a grande crise (ou *crash*) do mercado de telecomunicações iniciada com os anúncios de grandes prejuízos, seguidos da desvalorização e falência de várias empresas do setor em 2000/01.

de telecomunicações e as expectativas geradas pela chamada “revolução” da *Internet*.

Em primeiro lugar, a desregulamentação foi diretamente responsável pelo aparecimento de uma *bolha* de investimentos gigantesca. Embora a desregulamentação já tivesse começado em 1984 com a separação da AT&T em várias empresas nos Estados Unidos e a introdução de duas empresas concorrentes no Reino Unido, o *Telecom Act* de 1996 e a desregulamentação oficial do mercado pela União Européia em janeiro de 1998 representaram o mais importante processo de desregulamentação de qualquer setor da economia na história contemporânea. Outros países (inclusive o Brasil) adotaram, ao longo da década de 90, processos semelhantes de privatização e desregulamentação do setor de telecomunicações, fazendo com que tal processo de desregulamentação adquirisse proporções globais. Considerando que o setor de telecomunicações é de capital intensivo, o aumento de capital devido a entrada de novos investidores ansiosos por estabelecer sua presença nesse novo mercado foi responsável por uma enorme *bolha* de investimento de capital, que logo necessitaria de ser remunerada adequadamente.

Por outro lado, a disponibilização e o rápido crescimento da *Internet* utilizando a rede comutada (pública) de telefonia (*Public Switched Telephone Network* – PSTN) que foi inicialmente projetada e aperfeiçoada para transporte de voz, gerou um grande volume de tráfego IP que, a princípio, era inesperado e que se acreditava crescer “exponencialmente para sempre”. Esta “revolução” da *Internet* (em verdade, a expectativa de aumento da demanda devido ao crescente uso da *Internet*) forneceu mais incentivos para se gastar (investir capital) em novas tecnologias como, por exemplo, nas redes óticas e, em particular, em tecnologias como a WDM. Tais investimentos viriam não só aliviar de certa forma o congestionamento nas conexões existentes em uma PSTN mas também lançar as bases para o aparecimento de um novo tipo de rede no futuro: a rede ótica de transporte de alta capacidade. Contudo, naquele momento tais investimentos realizados principalmente nos *backbones* se mostraram incapazes de contornar os “gargalos” existentes nas redes de acesso, uma vez que a maioria dos usuários da *Internet* continuava utilizando as PSTNs para conexão com seus provedores de serviço.

O que se seguiu a isso foi o que normalmente acontece a todo e qualquer processo de desregulamentação em que a demanda tenha sido superestimada: um *crash*⁵[151]. Em resumo, como consequência da desregulamentação, muitos investidores novos entraram no mercado, reduzindo o preço para ganhar fatias do mesmo rapidamente e oferecendo produtos muito similares aos da concorrência. No curto prazo, a capacidade em excesso (principalmente, dos *backbones*) conduziu as empresas a um regime indisciplinado de descontos que juntamente com o congestionamento das redes de acesso (entre outros fatores) foram responsáveis pelo registro de enormes prejuízos, seguidos de uma desvalorização repentina de suas ações negociadas em bolsa, da retração dos investimentos no setor e, até mesmo, da falência de algumas delas.

Segundo os analistas do mercado de telecomunicações, no longo prazo, só sobreviverão aquelas empresas (quer provedores de serviços, quer fornecedores de equipamentos) que obtiverem ganhos com as economias de escala e que assegurem aumentos em sua produtividade. Desse modo, os primeiros anos deste século têm

⁵Este tipo de consequência foi primeiro observada e documentada em 1958 pelo economista George Stigler no artigo *The Economics of Scale* publicado pelo *Journal of Law and Economics*.

sido marcados por essa busca incessante por ganhos de produtividade. No caso específico do mercado de redes óticas, a preocupação fundamental não está mais relacionada à maximização da quantidade de comprimentos de onda que se é capaz de utilizar em um *backbone* e/ou equipamento ótico, mas sim com a quantidade mínima de comprimentos de onda que se pode realmente utilizar (para transportar todo o tráfego) ao menor custo operacional e de manutenção.

Consequentemente, as redes óticas estão deixando de ter um papel secundário (como uma tecnologia utilizada apenas para redução do congestionamento) para assumir uma posição central no contexto das futuras redes de telecomunicações. O tráfego continua crescendo⁶ – embora em um ritmo mais lento que o previsto pelas projeções do passado – e ainda há muita capacidade instalada. O que parece faltar são mecanismos eficientes e “inteligentes” de gerência/alocação dessa capacidade de forma a lidar com tal crescimento da demanda, ao mesmo tempo que se mantém o custo no menor patamar possível.

Para se alcançar esse objetivo, os analistas têm sido unânimes em apontar os seguintes aspectos como as principais tendências e linhas de pesquisas [151]:

- Integração das funcionalidades de comutação e transporte da camada ótica em um só componente de modo a reduzir a quantidade de equipamentos necessários aos elementos da rede;
- Utilização de tecnologia que reduza/elimine a necessidade de conversão E-O (eletro-ótica) para diminuir os atrasos decorrentes do processamento eletrônico, além de baratear os equipamentos;
- Adoção de esquemas de proteção e recuperação de falhas na camada ótica para redes irregulares, a fim de minimizar a capacidade reservada para proteção disponibilizando-a para a realização de outros serviços;
- Atualização dos equipamentos empregados nas redes metropolitanas visando permitir serviços de banda larga em diferentes tipos de arquitetura de rede e diferentes protocolos (em especial, para os protocolos IP e Ethernet);
- Utilização de *transceivers* (transmissores e receptores óticos) “ajustáveis” para minimizar o custo e tempo de reparo de um equipamento WDM; e
- Implementação e uso de comutação “híbrida” através da adoção de uma matriz de comutação ótica e outra eletro-eletrônica para que tal equipamento “híbrido” possa eficientemente comutar comprimentos de onda ao mesmo tempo em que realiza *grooming* de tráfegos de baixa velocidade.

Como discutido anteriormente, todas essas tendências e linhas de pesquisa e desenvolvimento constituem a base para o futuro pós-*crash* e, portanto, têm em comum uma única meta: reduzir os custos operacionais e de manutenção das empresas

⁶Segundo [134], o tráfego da *Internet* cresceu 116% na Europa em 2002, alcançando a marca de 95 Pb/mês. Isto representa pouco mais da metade do volume nos EUA, onde o crescimento em 2002 foi de 85%. Já as estimativas de retorno do investimento totalizaram em 2002 para a Europa e para os EUA, respectivamente, US\$ 8,8 bilhões e US\$ 14 bilhões. Isto representa uma redução no retorno do investimento por bit em relação a 2001 de 48%, na Europa, e 52%, nos EUA.

Tabela 1.1: Matriz de Tráfego do Exemplo

Elemento de Origem	Elemento de Destino					
	1	2	3	4	5	6
1	0	0	0	125	25	0
2	0	0	0	50	25	0
3	0	0	0	0	0	50
4	400	100	0	0	0	50
5	100	300	0	0	0	0
6	0	0	50	50	0	0

“sobreviventes”. No tocante a este trabalho, deve-se destacar seu relacionamento estreito com o último tópico da lista acima.

Cumpra, ainda, dizer que o futuro das redes óticas continua sendo muito promissor; contudo, o mercado global jamais deverá alcançar novamente a marca histórica de US\$ 30 bilhões atingida em 2000 devido principalmente ao aumento da integração⁷, uma vez que um elemento de rede do futuro será capaz de realizar o que era feito por 20 a 30 elementos em 2000 [151]. Vale destacar que os setores ligados aos aspectos mencionados acima têm experimentado uma menor retração. Em 2002, o setor de novos equipamentos “híbridos” capazes de realizar comutação ótica e *grooming* sofreu uma redução em relação a 2001 muito menor (de apenas 17% na Europa [137], por exemplo) do que a ocorrida nos demais fornecedores de equipamentos (a redução do mercado global de equipamentos para redes óticas chegou a 65%, em 2002 [136]).

Por fim, antes de se apresentar formalmente os objetivos deste trabalho, procurar-se-á demonstrar através de um exemplo prático os ganhos obtidos no projeto de uma rede ótica a partir da obtenção de uma “boa” solução para o **TGP**.

1.1.3 Um Exemplo Prático

A fim de ilustrar de forma prática os diversos fatores (e ganhos) envolvidos em uma situação em que se utiliza do *Traffic Grooming*, considere o seguinte exemplo adaptado de [45]. Seja a rede em anel com seis elementos da Figura 1.1(a) uma representação da topologia de um anel ótico unidirecional. Considere ainda que a Tabela 1.1 fornece o volume de tráfego (em canais de 2 Mbps) que deve ser transportado entre os elementos da rede e que a capacidade de transporte de um comprimento de onda é igual a 500 canais (≈ 1 Gbps).

A Figura 1.1(b) exhibe uma forma possível de se transportar todo o tráfego através do anel, em que se utiliza 2 comprimentos de onda que são inseridos e retirados em cada elemento da rede. Neste caso os elementos são equipados apenas com comutadores/roteadores eletro-eletrônicos e devem realizar a conversão ótico-eletrônica (OEO) do tráfego. Esta solução não necessita de roteamento ótico dos comprimentos de onda (*wavelength routing*), contudo, requer uma significativa capacidade de

⁷Segundo [135, 138], as atuais estimativas prevêem um crescimento de 8% no mercado global de redes óticas. Isto equivale a dizer que somente em 2006 o mercado deverá retornar ao patamar em que se encontrava em 1998.

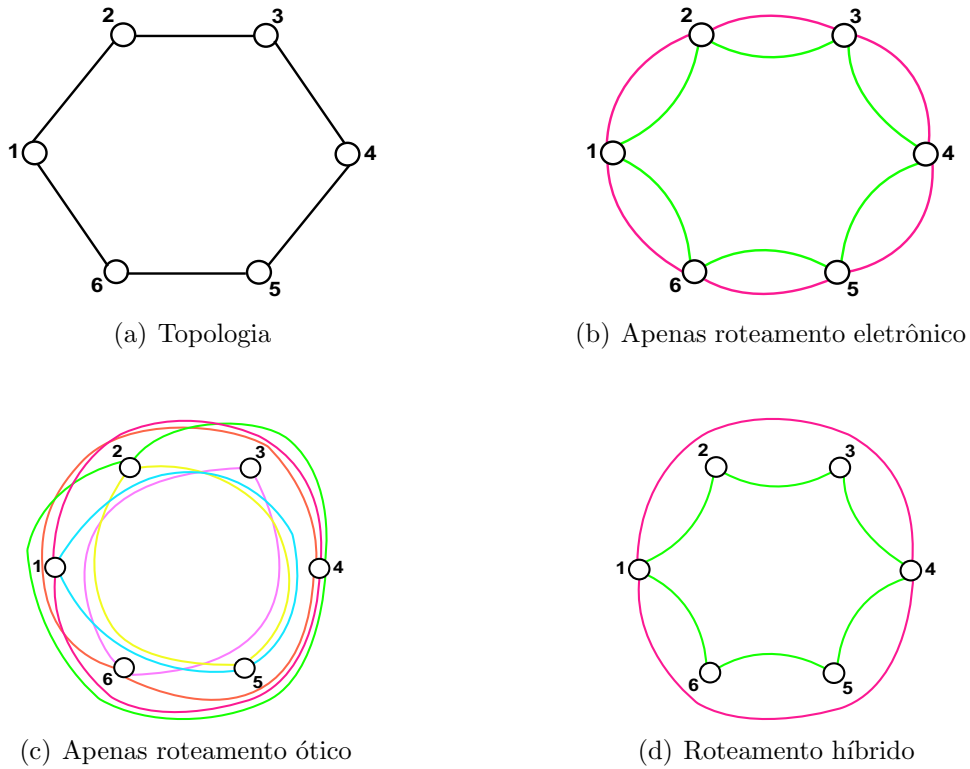


Figura 1.1: Exemplo Prático de *Traffic Grooming*

comutação/roteamento eletro-eletrônico em cada elemento do anel de modo a rotear todo o tráfego para o seu destino. A quantidade total de tráfego comutado/roteado eletronicamente em todos os elementos da rede é da ordem de 5250 Mbps (2625 canais de 2 Mbps), que representa o somatório da quantidade de tráfego que passa através de cada um dos elementos da rede (mas que não se origina ou se destina a ele). Este valor representa quase o dobro do volume total de tráfego na matriz. Em outras palavras, em média, cada componente de tráfego (canal) é comutado/roteado em dois elementos intermediários antes de alcançar seu destino. Além disso, esta solução necessita de 12 ADMs, uma vez que cada elemento do anel deve retirar e inserir os dois comprimentos de onda utilizados. Uma vez que o *grooming* geralmente se refere a utilização apropriada de um esquema híbrido (ótico e eletrônico) de roteamento, pode-se referir a essa solução como sendo uma em que nenhum *grooming* foi realizado.

Uma outra solução em que não se realiza nenhum *grooming* é exibida na Figura 1.1(c), só que desta vez ela representa o extremo oposto. Neste caso, cada componente de tráfego é transportado através de um *lightpath* ligando diretamente sua origem a seu destino. Cada uma dos comprimentos de onda é roteado opticamente (*wavelength routing*) em cada elemento intermediário (por exemplo, através do uso de OADMs, ver subseção 2.3.2) e nenhuma comutação/roteamento eletrônico é realizado a não ser nos elementos de origem e destino. Todavia, esta solução necessita de um grande número de comprimentos de onda, equivalente a número máximo de componentes distintos (cuja origem e destino são diferentes) de tráfego que devem ser transportados através de um enlace físico (neste caso, igual a seis).

Em uma situação real, essa quantidade elevada de comprimentos de onda pode não ser disponível e, caso seja, esta abordagem resultaria em uma subutilização severa dos recursos da rede. Além disso, esta solução não reduz a quantidade de LTEs necessários, visto que o número de ADMs utilizados é o mesmo da solução anterior. Este resultado indica que a completa eliminação do roteamento eletrônico não contribui necessariamente para uma redução no custo dos componentes utilizados no projeto da rede.

Finalmente, a Figura 1.1(d) exibe uma solução entre os dois cenários extremos apresentados anteriormente, em que se utiliza uma combinação apropriada de roteamento ótico e roteamento eletrônico visando reduzir o custo total da rede ao mesmo tempo em que se utiliza apenas 2 comprimentos de onda. Um dos comprimentos de onda é retirado e inserido (e roteado eletronicamente) em cada elemento da rede e todo o tráfego entre os elementos 2, 3, 5 e 6 é transportado através dele, bem como o tráfego originado em 4 com destino ao elemento 6 e os tráfegos originados em 2 e 6 destinados ao elemento 4. Já o tráfego do elemento 1 para os demais elementos e o tráfego do elemento 4 para os elementos 1 e 2 são transportados através do segundo comprimento de onda. Este comprimento de onda só é retirado e inserido da rede nos elementos 1 e 4 (sendo roteado óticamente nos demais elementos), de modo que alguns componentes do tráfego (por exemplo, o tráfego de 4 para 2) devem ser roteados eletronicamente para o primeiro comprimento de onda em algum elemento intermediário (para o exemplo dado, no elemento 1). Esta solução necessita de somente 8 ADMs. Além disso, ela também atinge o valor mínimo possível de roteamento eletrônico utilizando não mais de 2 comprimentos de onda, sendo neste sentido ótima [45].

Este exemplo demonstra claramente que a realização adequada de *grooming* pode se traduzir em benefícios significativos em termos de redução do custo da rede (associado diretamente ao número de ADMs) e/ou número de comprimentos de onda necessários em comparação com as soluções extremas apresentadas em que se utiliza roteamento eletrônico ponto-a-ponto ou roteamento completamente ótico.

1.1.4 Sobre um Enfoque de Otimização

A despeito da obtenção de algum sucesso na resolução do **TGP** para redes com topologia física em anel, tais abordagens (em sua maioria, através de heurísticas) não se adaptam facilmente para os casos em que a rede possui uma topologia diversa (por exemplo, irregular, anéis interconectados ou cadeia de anéis). Dessa forma, novos esforços de pesquisa têm sido realizados no sentido de se solucionar adequadamente o problema para redes com topologia física irregular ou para um conjunto de anéis interconectados.

Caso o enfoque dado seja solucionar o **TGP** em uma rede cuja topologia física é arbitrária, o número de trabalhos encontrados na literatura se reduz a poucos. Independentemente disto, observa-se em [45] que tais trabalhos têm sido marcados por um carácter eminentemente experimental em que, salvo raras exceções, pouca ou nenhuma análise “mais formal” dos resultados foi realizada, ou ainda, nenhuma medida de desempenho (qualidade dos resultados) foi caracterizada formalmente como, por exemplo, através de limites inferiores e superiores “bem justos” (isto é,

próximos da solução ótima). De fato, algumas métricas têm sido usadas na avaliação desses resultados (tais como o número mínimo e máximo de comprimento de ondas e/ou ADMs necessários). Contudo, valores mais precisos ou “justos” para tais limites não são conhecidos, salvo raras exceções⁸.

Nesse sentido, pode-se afirmar que é necessário o desenvolvimento de abordagens sistemáticas e formais para a resolução do **TGP** cujos desempenhos (qualidade dos resultados obtidos) possam ser caracterizados e avaliados não só para o preenchimento de uma lacuna teórica, mas principalmente devido à natureza eminentemente prática e operacional desse problema, cuja resolução se faz necessária a todo e qualquer projeto de rede ótica.

Os autores em [45] chegam a sugerir a adoção de técnicas da Teoria dos Grafos juntamente com algoritmos cujas propriedades possam ser (ou já tenham sido) formalmente verificadas, de modo a se desenvolver algoritmos de resolução que sejam bastante flexíveis e eficientes para lidar com diferentes tipos de topologias e que, ao mesmo tempo, apresentem propriedades formais passíveis de verificação.

Este trabalho procura atender a esses requisitos propondo para isso abordagens sistemáticas cuja qualidade dos resultados possa ser formalmente caracterizada e mensurada. Posto isto, apresenta-se na seção seguinte os objetivos deste trabalho.

1.2 Objetivos

Como exposto anteriormente, o objetivo deste trabalho é investigar o **TGP** (juntamente com algumas de suas variações) independentemente da topologia da rede física subjacente. Para tanto duas abordagens distintas são adotadas: um algoritmo de *branch-and-bound* baseado no uso da relaxação lagrangeana e um algoritmo de *branch-and-cut* utilizando inequações válidas do poliedro associado ao conjunto de soluções do problema. Um detalhamento dessas abordagens é realizado mais adiante.

Antes, porém, tomando-se por base uma formulação natural do problema, uma representação da rede através de um grafo em camadas é proposta e utilizada no desenvolvimento de uma “nova” formulação simplificada para o mesmo que permita constatar o relacionamento existente entre o **TGP** e outros problemas já estudados na literatura. Essa representação em camadas apresenta uma série de propriedades e características que são estudadas e exploradas nas duas abordagens de resolução. Além disso, tal formulação simplificada será utilizada como base para o desenvolvimento dos diversos algoritmos.

No tocante às variações do **TGP** mencionadas acima, além do problema de *Traffic Grooming* genérico descrito anteriormente, este trabalho aborda duas variações do mesmo. Na primeira delas, é investigada uma versão do problema em que a realização de *grooming* (combinação de tráfegos de baixa velocidade em canais de alta velocidade) está limitada a um subconjunto dos elementos da rede. Esta versão do problema possui um apelo eminentemente prático uma vez que o custo dos equipamentos necessários para realização de *grooming* pode inviabilizar sua disponibilidade por toda a rede (principalmente para aquelas com um grande número

⁸Limites “justos” são conhecidos apenas para as situações em que a rede possui uma topologia bem simples (como no caso dos anéis unidirecionais) e o padrão de tráfego é bem “comportado”, ou melhor, quando se utiliza de hipóteses bem restritivas (como supor que o tráfego é uniforme).

de elementos). Já na segunda variação, o problema de *Traffic Grooming* em uma rede sujeita a “reconfigurações” ao longo do tempo será investigado. A possibilidade de reconfiguração de uma rede ótica WDM (inclusive sem a interrupção das conexões já existentes) é uma de suas grandes vantagens. Em verdade, esses problemas são bastante complexos, mesmo que não se considere o *Traffic Grooming*. Dessa forma, apenas uma investigação preliminar foi realizada, procurando determinar se as abordagens desenvolvidas para o problema genérico podem ou não ser aplicadas com sucesso a essas situações.

Tudo o que foi exposto anteriormente pode ser resumido da maneira que se segue em que se estabelece o tema deste trabalho e suas delimitações, bem como seu objetivo geral e demais objetivos secundários.

1.2.1 Tema

Traffic Grooming em Redes Óticas WDM.

1.2.2 Delimitação do Tema

Estudo sobre *Traffic Grooming* em redes óticas WDM independentemente da topologia da rede física subjacente. Além disso, será realizada uma investigação sobre a adequação dos métodos desenvolvidos na resolução de uma versão do problema em que se considere a reconfiguração da rede ao longo de um horizonte de tempo limitado e de outra, em que apenas alguns dos elementos da rede são capazes de realizar *grooming*.

1.2.3 Objetivo Geral

O objetivo geral deste trabalho é desenvolver algoritmos de resolução para o **TGP** bem como testar o desempenho (qualidade dos resultados obtidos) dos mesmos.

1.2.4 Objetivos Específicos

Entre os objetivos específicos do desenvolvimento deste trabalho, pode-se citar os seguintes:

- Propor uma representação da rede através de um grafo em camadas que permita obter uma formulação simplificada para o **TGP**, além de permitir relacioná-lo a outros problemas da literatura;
- Obter através do uso da relaxação lagrangeana um conjunto de limites inferiores para o **TGP**;
- Desenvolver e testar algoritmos para maximizar o valor dos limites inferiores obtidos através da relaxação lagrangeana;
- Desenvolver e testar heurísticas para o cálculo de limites superiores para o **TGP**;

- Desenvolver e testar um algoritmo de *branch-and-bound* baseado no uso de limites inferiores obtidos através da relaxação lagrangeana;
- Estudar a estrutura facial do poliedro associado ao conjunto de soluções do problema para se obter um conjunto de inequações válidas (e, eventualmente, facetas) para o **TGP**;
- Desenvolver e testar um método de *planos-de-corte* utilizando os resultados do estudo da estrutura facial do poliedro associado ao conjunto de soluções do **TGP**;
- Desenvolver e testar um algoritmo de *branch-and-cut* utilizando inequações válidas do poliedro associado ao conjunto de soluções do problema.

Além disso, deve-se destacar que, durante o desenvolvimento e testes tanto da abordagem lagrangeana quanto da abordagem poliédrica, alguns resultados adicionais foram obtidos merecendo destacar as implementações e testes da relaxação linear e de um algoritmo de *cut-and-branch*⁹ para o **TGP**.

1.3 Principais Contribuições

A partir dos objetivos apresentados na seção anterior, cumpre destacar as principais contribuições geradas pela realização do presente trabalho. Sendo assim, pode-se citar as seguintes contribuições:

- Proposição de uma representação estendida para a topologia de rede a partir da qual uma “nova” formulação natural para o **TGP** é apresentada (essa formulação é capaz de capturar os detalhes tecnológicos envolvidos no problema evitando, contudo, redundâncias desnecessárias)¹⁰;
- Proposição de uma representação de um grafo em camadas para a topologia de rede cuja adoção permite uma simplificação da formulação apresentada para o **TGP**, possibilitando, assim, relacioná-lo a outros problemas estudados na literatura;
- Desenvolvimento de algoritmos exatos (e heurísticas) para resolução do **TGP** e algumas de suas variações;
- Estudo da estrutura facial do poliedro associado ao conjunto de soluções do **TGP**;
- Análise comparativa da qualidade dos limites inferiores produzidos através da relaxação lagrangeana e da abordagem poliédrica (método de *planos-de-corte*);

⁹Em um algoritmo de *cut-and-branch*, após a aplicação de um método de *planos-de-corte* ao problema original, utiliza-se de um algoritmo de *branch-and-bound* “comum” na busca da solução ótima, isto é, durante a busca na árvore de subproblemas não há nenhuma tentativa de se gerar novos *planos-de-corte*.

¹⁰As formulações para o **TGP** existentes na literatura, quando comparadas àquela apresentada neste trabalho, não são suficientemente gerais ou são demasiadamente redundantes para serem consideradas “eficientes”, como será discutido mais adiante (ver subseção 2.4).

- Análise comparativa dos métodos exatos (e heurísticas) desenvolvidos para resolução do **TGP** com algoritmos apresentados na literatura.

1.4 Organização do Texto

O texto deste trabalho se encontra organizado da seguinte forma. No capítulo seguinte, apresenta-se uma revisão da literatura relacionada ao **TGP**. O capítulo 3 introduz uma representação em camadas da rede a partir da qual se desenvolve um conjunto de formulações de programação matemática para o **TGP** (e algumas de suas variações) que são utilizadas no restante deste trabalho. As abordagens propostas para obtenção de métodos de resolução do **TGP** descritas no capítulo 4. Já os capítulos 5 e 6 descrevem para cada uma das abordagens – lagrangeana e poliédrica – os detalhes relativos a sua formalização, bem como os resultados dos experimentos computacionais. Uma análise comparativa entre os diversos métodos propostos pelas duas abordagens é apresentada no capítulo 7. O capítulo 8, por sua vez, discute a adequação das abordagens propostas para o **TGP** na resolução de versões modificadas do problema original – **TGRP** e **STGP**. O último capítulo apresenta as conclusões obtidas durante a realização deste trabalho, juntamente com algumas propostas para trabalhos futuros.

Capítulo 2

Revisão da Literatura

SEGUNDO os autores de [45], o **TGP** é essencialmente uma variante do problema de projeto de uma topologia virtual para uma rede ótica. Entretanto, a maioria dos trabalhos sobre projeto de topologias virtuais não enfoca o *Traffic Grooming* nem os aspectos de minimização de custo relacionados a ele. Em verdade, tais estudos consideram que o conjunto de *lightpaths* a serem estabelecidos é dado inicialmente ou que uma taxa média de tráfego entre todos os pares origem-destino é conhecida, enquanto que em relação à função objetivo as métricas consideradas de interesse são o número de comprimentos de onda ou o nível de congestionamento da rede. Consequentemente, estas abordagens não podem ser aplicadas diretamente em contextos como o das redes óticas SONET/SDH com topologia em anel, em que o objetivo é combinar (*groom*) um conjunto de componentes de tráfego (conexões) independentes e discretos de modo a minimizar o custo de rede ou o roteamento eletrônico.

As seções seguintes são dedicadas a uma revisão da literatura relacionada ao **TGP**. Em primeiro lugar se fará uma revisão da literatura sobre o projeto de redes óticas e, em especial, sobre o projeto de topologias virtuais. Em seguida, o principal problema relacionado ao projeto de topologias virtuais, o problema de roteamento e atribuição de comprimentos de onda (*Routing and Wavelength Assignment – RWA*) é discutido. Por fim, a última seção apresenta uma revisão da literatura sobre o **TGP** propriamente dito procurando destacar os diferentes trabalhos de pesquisa (e seus respectivos resultados) para redes óticas com topologia em anel e irregular.

2.1 Sobre Projeto de Redes Óticas

O problema de planejamento e projeto de redes óticas pode ser visto como uma variação do problema equivalente em uma rede multifluxo em que cada fluxo de produto (ou *commodity*) passa a representar uma conexão para fluxo de dados. Os primeiros esforços – ainda que isolados e limitados – no sentido de se abordar de maneira mais sistemática esse problema datam do início da década de 90. Contudo, tais trabalhos – como destacado mais adiante – em sua grande maioria se limitam a discussões em torno das características físicas das redes e em como dispor (localizar)

seus principais elementos da melhor forma, principalmente em virtude das restrições tecnológicas e de custo ainda existentes nessa época.

Abordagens mais sistemáticas e profundas somente surgem na literatura a partir de meados da década de 90 e, principalmente, em seu final, em razão do novo impulso dado nas pesquisas em busca de novos meios e técnicas de alta capacidade e escalabilidade para o transporte de dados, por razões já mencionadas anteriormente.

Tradicionalmente, o problema de planejamento e projeto de redes óticas tem sido dividido em dois grandes subproblemas: o projeto da rede física a ser implantada e o projeto de uma topologia virtual (rede virtual) que ficará encarregada em última instância do processo de comunicação (transporte de dados) entre pontos de oferta e demanda, utilizando-se para tanto da rede física projetada anteriormente.

O primeiro dos subproblemas citados procura estabelecer quantos e quais elementos (repetidores/amplificadores, *switches*, conversores, entre outros) irão interligar os pontos de oferta e demanda, além de determinar quais rotas físicas (caminhos através de uma ou mais fibras óticas) irão existir entre tais elementos. Uma descrição mais detalhada dos componentes de uma rede ótica pode ser encontrada em [23, 133]. Dado que na maioria dos casos as redes físicas já existem (normalmente redes de longa distância ou em grandes áreas metropolitanas) e que foram implementadas ao longo dos anos na medida da necessidade (segundo critérios *ad-hoc* e heurísticas), nenhum esforço, exceto o trabalho em [36], no sentido de solucionar este subproblema por completo foi encontrado na literatura até o presente momento. De fato, muitos dos trabalhos que procuraram tratar o projeto da rede física se limitaram apenas ao problema de localização dos amplificadores em uma rede já existente, procurando minimizar dessa forma a quantidade necessária dos mesmos e, conseqüentemente, os custos de sua implantação.

Mais recentemente, tal subproblema recebeu novas contribuições, devendo-se mencionar o trabalho em [77] pelo estudo dos problemas decorrentes do uso de amplificadores em *loops* (anéis fechados de fibras óticas), além de uma investigação sobre a utilização de conversores de comprimento de onda de capacidade limitada (isto é, restrita apenas a alguns dos comprimentos de onda disponíveis/utilizados na rede), o que tende a reduzir o custo de tais equipamentos que podem ser essenciais se for aplicado o roteamento através do comprimento de onda. Em verdade, no trabalho apresentado em [131] são definidos pela primeira vez alguns dos termos relacionados à conversão limitada de comprimento de onda; além disso, os autores foram capazes de estabelecer limites teóricos sobre o número de conversores necessários na implementação de determinadas topologias. Ainda sobre o mesmo subproblema, deve-se destacar o trabalho em [130] que procurou determinar uma localização ótima de amplificadores para redes óticas locais e metropolitanas que se utilizam de mecanismos de *broadcasting*.

Quanto ao subproblema de projeto da topologia virtual (rede virtual), os primeiros estudos também datam do início da década de 90, com destaque para os trabalhos apresentados em [8, 31, 32, 105]. Contudo, diferentemente do ocorrido com o subproblema anterior, a partir de meados da década 90 – época em que muitas inovações tecnológicas passaram a ser investigadas e, conseqüentemente, noticiadas – pode-se encontrar na literatura inúmeros trabalhos abordando em parte ou integralmente o projeto da rede virtual. Segundo os autores de [43], esse problema pode ser decom-

posto, por sua vez, de maneira aproximada em quatro subproblemas distintos. Tal decomposição é aproximada devido ao fato de que a resolução de tais subproblemas em sequência e a combinação de suas soluções pode não resultar na obtenção da solução ótima para o problema original de projeto da rede virtual ou, até mesmo, pode se chegar a uma situação em que nenhuma solução é obtida. De todo modo, essa decomposição, proposta originalmente em [113], é consistente com a abordagem feita em vários trabalhos na literatura, devendo-se destacar entre eles os trabalhos apresentados em [6, 91, 112, 132]. Por fim, a decomposição mencionada consiste em subdividir o problema de projeto da rede virtual em 4 subproblemas: determinação da topologia lógica (ou virtual), roteamento de canais lógicos de comunicação (*lightpaths*), atribuição de comprimentos de onda e, finalmente, roteamento do tráfego. Cada um desses subproblemas será detalhado a seguir.

O subproblema de determinação da topologia lógica (ou virtual) procura estabelecer qual topologia deverá ser imposta sobre a rede física independentemente da topologia desta última, de modo a estabelecer quais serão os canais lógicos de comunicação (*lightpaths*) em termos dos pontos de oferta e demanda da rede. Já o subproblema de roteamento de canais lógicos de comunicação (*lightpaths*) busca determinar quais ligações físicas (fibras) serão utilizadas por cada um dos canais lógicos estabelecidos anteriormente, ou melhor, qual a rota seguida por cada canal através da rede física. O subproblema de atribuição de comprimentos de onda, por sua vez, determina qual comprimento de onda é utilizado por cada canal lógico (*lightpath*) da topologia virtual levando-se em consideração as restrições capacidade e conversão existentes nas ligações físicas (fibras da rede física). Finalmente, o subproblema de roteamento do tráfego cuida da efetiva transmissão dos produtos (dados) entre os pontos de oferta e demanda através da topologia virtual obtida (isto é, definição dos fluxos através das rotas já previamente estabelecidas).

Como já mencionado anteriormente, a abordagem padrão para esses problemas até o presente momento tem sido o desenvolvimento de heurísticas. Tais heurísticas encontradas na literatura podem ser agrupadas em três diferentes categorias [43]. A primeira delas agrupa os trabalhos em que a topologia virtual a ser implementada é na verdade uma topologia regular bem difundida e estudada, tal como um hipercubo, uma estrela ou anéis multiconectados. O uso de uma topologia regular amplamente conhecida traz consigo várias vantagens pelo fato de serem bem compreendidas, do roteamento do tráfego ser facilmente implementado e de se encontrar bastante literatura a seu respeito. O trabalho em [9] apresenta uma excelente avaliação de diversas topologias lógicas para uma rede ótica abordando, além de uma descrição de seu funcionamento e suas principais propriedades, detalhes sobre roteamento, tolerância a falhas e escalabilidade. Nesta primeira categoria pode-se citar as heurísticas propostas nos trabalhos apresentados em [32, 105, 112, 113].

Em uma segunda categoria encontram-se os trabalhos em que os canais lógicos de comunicação (*lightpaths*) da topologia virtual já são conhecidos (fornecidos como dados de entrada) em termos de suas origens e destinos. Neste contexto a maioria das heurísticas enfoca o subproblema de roteamento de canais lógicos de comunicação (*lightpaths*) e, eventualmente, os subproblemas de atribuição de comprimentos de onda e roteamento do tráfego também são tratados. Nesta categoria existem inúmeros trabalhos, merecendo destaque aqueles apresentados em [6, 28, 31].

Finalmente, na terceira e última categoria estão as heurísticas e trabalhos em que nenhuma suposição sobre a topologia virtual é feita inicialmente. Esta categoria tem atraído um maior interesse recentemente, uma vez que estes estudos abordam diretamente o subproblema de determinação da topologia lógica, além de explorarem alguns dos outros subproblemas do projeto da rede virtual (eventualmente até todos). Destacam-se nesta linha de atuação os trabalhos em [7, 10, 91, 132, 165].

Vários trabalhos na literatura têm abordado de forma conjunta os subproblemas de roteamento de *lightpaths* e atribuição de comprimentos de onda (*Routing and Wavelength Assignment – RWA*), que são considerados os problemas fundamentais na obtenção de uma topologia virtual¹. A próxima seção é dedicada a discussão de tais trabalhos e de seus principais resultados.

Um outro problema intimamente relacionado com o problema de projeto da rede virtual é a reconfiguração da rede ótica de uma topologia virtual para outra. Uma primeira abordagem assume que tanto a topologia atual quanto a nova são conhecidas a priori e o problema se resume a minimizar o custo relativo às alterações que se fazem necessárias para se obter a nova topologia a partir da atual. Apesar de não terem sido encontrados estudos dessa natureza para redes óticas de grandes distâncias, pode-se encontrar em [92] um levantamento detalhado da literatura a esse respeito para redes locais. Uma outra abordagem desse problema pode ser vista no trabalho em [7] em que se assume que apenas a topologia atual é fornecida, juntamente com o padrão de modificação do tráfego e/ou da topologia física (tais alterações tornam a reconfiguração extremamente necessária).

Por fim, convém comentar o trabalho em [139] em que o conceito de canal ou caminho lógico de comunicação (*lightpath*) entre dois pontos (um de oferta e outro de demanda) é generalizado de modo a possuir vários pontos de destino (isto é, um nó de oferta ligado a vários de demanda). Dessa forma surge o conceito de *lighttree*. Um dos pontos fundamentais desse trabalho é a demonstração de que, dado um problema de projeto de rede virtual, uma solução ótima em que se utilize de *lighttrees* é tão boa quanto aquela que seria obtida através do uso de *lightpaths* (eventualmente até melhor). Outro fator importante, que não havia sido abordado até então, reside no fato de que os *lighttrees* permitem a representação (e implementação) do *multicasting* ótico – que pode vir a ser necessário para alguns dos serviços prestados no futuro, como HDTV.

2.2 Sobre o *RWA*

Uma rede ótica pode ser representada por um grafo $H = (V, E)$, em que cada nó representa um elemento da rede e cada arco representa um enlace físico (fibra ótica) entre dois elementos da mesma. Uma requisição r de conexão entre os nós u e v é satisfeita através da (i) atribuição a r de um caminho p_r entre u e v em H , e (ii) atribuição a p_r de um comprimento de onda para transportar a informação através de seus arcos (o caminho p_r e o comprimento de onda associado constituem um *lightpath*). Seja R o conjunto de todas as requisições de conexão (em verdade, R

¹Os demais subproblemas são considerados de menor importância e ignorados completamente em tais estudos.

é um multiconjunto em que um elemento pode ocorrer várias vezes). Sendo assim, o problema de *RWA* procura satisfazer todas as requisições em R de modo que se duas requisições r e r' possuírem caminhos p_r e $p_{r'}$, respectivamente, e esse caminhos compartilharem algum arco, então diferentes comprimentos de onda são atribuídos a eles. O objetivo é satisfazer todas as requisições em R utilizando o número mínimo de comprimentos de onda possível. O problema de *RWA* não envolve roteamento eletrônico (apenas roteamento ótico) e, geralmente, o mesmo comprimento de onda é atribuído a todos os arcos ao longo do caminho p_r da requisição r (fato conhecido, na literatura, como restrição de *continuidade* de comprimento de onda).

Em primeiro lugar, deve-se notar que se H for uma árvore, então cada par de elementos é conectado por um único caminho e a primeira parte do problema de *RWA* está implicitamente resolvida. Além disso, segundo [45], o problema de atribuição de comprimentos de onda de modo a minimizar o número de comprimentos de onda utilizado pode ser resolvido em tempo polinomial para redes cuja topologia é em cadeia, em estrela ou *spider* (uma rede com topologia *spider* é uma árvore com exatamente um vértice de grau maior que dois [159]), entretanto ele é \mathcal{NP} -difícil para árvores em geral.

Se H é uma rede com topologia em cadeia, o problema de atribuição é equivalente ao problema de coloração em um grafo intervalar, que, por sua vez, pode ser resolvido em tempo linear utilizando um algoritmo guloso [68]. Caso H seja uma rede com topologia em estrela, o problema de atribuição equivale a encontrar uma coloração mínima para as arestas em um grafo bipartite, que pode ser resolvido em tempo polinomial utilizando os teoremas de Hall e König [158]. Já a minimização do número de comprimentos de onda para redes com topologia *spider* também é possível em tempo polinomial [159].

Contudo se H for uma árvore qualquer, o problema de minimização do número de comprimentos de onda é \mathcal{NP} -difícil, mesmo se H for uma árvore binária [35]. Apesar disso, dado um conjunto qualquer de requisições em uma rede em árvore, existe um algoritmo polinomial para gerar uma atribuição de comprimentos de onda que utiliza não mais que $5/2$ do valor ótimo [84].

No caso de uma rede com topologia em anel, uma requisição pode ser roteada de duas formas distintas, fazendo que ambos os problemas (i) e (ii) devam ser considerados pelo algoritmo de resolução. Este problema é \mathcal{NP} -difícil [159]. Contudo, para um conjunto qualquer de requisições em uma rede em anel, é possível se obter em tempo polinomial um conjunto de caminhos juntamente com os respectivos comprimentos de onda para rotear as requisições, de modo que o número de comprimentos de onda utilizado não seja maior que o dobro do valor ótimo [159].

O problema de *RWA* é \mathcal{NP} -difícil em topologias arbitrárias e permanece assim mesmo em casos especiais como, por exemplo, quando cada *lightpath* for limitado a no máximo dois arcos (enlaces físicos) [69]. Entretanto, o problema de *RWA* se torna mais fácil quando algum grau de conversão de comprimento de onda é permitido nos OXCs (ou em alguns deles). Por exemplo, sabe-se que, no caso de apenas um único elemento ser capaz de conversão plena de comprimentos de onda em uma rede com topologia em anel, o número de comprimentos de onda necessários é igual ao número máximo de *lightpaths* que passam por qualquer arco da rede [131]. Mais especificamente, o trabalho em [131] foi capaz de demonstrar que mesmo uma

capacidade de conversão bastante limitada pode promover ganhos significativos na capacidade de transporte de redes com topologia em anel, em estrela e em árvore; e, até mesmo, para topologias arbitrárias. Por fim, resultados recentes demonstram melhorias similares para redes multifibra [98, 104] (vale dizer que um rede multifibra é equivalente, sob a ótica do *RWA*, a uma rede em que cada OXC é capaz de realizar conversões limitadas de comprimentos de onda, situação similar ao modelo apresentado em [131]).

2.3 Sobre o TGP

Pode-se formalizar o **TGP** de maneira análoga ao *RWA*. Como antes, uma rede ótica pode ser representada por um grafo $H = (V, E)$, em que cada nó representa um elemento da rede e cada arco representa um enlace físico (fibra ótica) entre dois elementos da mesma. Contudo, cada requisição r está associada a um valor inteiro c_r que representa a demanda (isto é, quantidade de capacidade necessária para se transportar o respectivo tráfego). Além disso, cada comprimento de onda possui uma capacidade máxima de transporte igual a C . Uma requisição r de conexão entre os nós u e v é satisfeita através da (i) atribuição a r de um caminho p_r entre u e v em H , e (ii) atribuição a p_r de um comprimento de onda para transportar a informação através de seus arcos (o caminho p_r e o comprimento de onda associado constituem um *lightpath*). Seja R o conjunto de todas as requisições de conexão (como antes, R é um multiconjunto em que um elemento pode ocorrer várias vezes). Sendo assim, o problema de **TGP** procura satisfazer todas as requisições em R de modo que dadas duas requisições r e r' que possuem caminhos p_r e $p_{r'}$, respectivamente, se um deles estiver completamente contido no outro e $c_r + c_{r'} \leq C$ então o mesmo comprimento de onda pode ser atribuído a ambos; caso contrário, diferentes comprimentos de onda devem ser atribuídos a eles. O objetivo é satisfazer todas as requisições em R utilizando o número mínimo de caminhos possível.

2.3.1 Técnicas de Multiplexação

Multiplexação por divisão de comprimento de onda (WDM) é uma tecnologia promissora no tocante a acomodação do crescimento explosivo da Internet e do tráfego de telecomunicação em redes de longa distância e metropolitanas e, até mesmo, em redes locais. Uma única fibra ótica possui uma largura de banda potencial de 50 THz. Utilizando-se da WDM, esta largura de banda pode ser dividida em múltiplos canais com frequência ou comprimento de onda não sobrepostos. Cada canal WDM pode ser operado, pelo menos a princípio, a “qualquer velocidade”, por exemplo, a velocidade de pico do processamento eletrônico em torno de alguns gigabits por segundo [111, 133]. Atualmente, fibras óticas comercialmente disponíveis podem suportar cerca de uma centena de canais (comprimentos de onda), podendo cada um deles operar a uma velocidade de transmissão acima de um gigabit por segundo (por exemplo, 2,5 Gbps, 10 Gbps e, em um futuro próximo, 40 Gbps).

Enquanto uma única fibra possui algumas dezenas de terabits por segundo de largura de banda [110] e um canal de comprimento de onda tem mais de um gigabit por segundo de velocidade de transmissão, as redes, via de regra, precisam

transportar tráfego a taxas muito mais baixas que tais capacidades. A exigência de capacidade desses tráfegos de baixa velocidade podem variar desde de um STS-1 (51.84 Mbps ou menos) até toda a capacidade de um comprimento de onda. Sendo assim, de modo a se reduzir o custo da rede e melhorar seu desempenho, é muito importante que o operador da rede seja capaz de combinar múltiplos tráfegos de baixa velocidade em canais de alta capacidade. *Grooming* é o termo utilizado para descrever a otimização da utilização da capacidade em sistemas de transporte [14].

Para tanto, diferentes técnicas de multiplexação podem ser usadas nas redes óticas WDM como, por exemplo:

- Multiplexação por divisão de espaço (SDM) – particiona o espaço físico para aumentar largura de banda, por exemplo, agrupando um conjunto de fibras em um único cabo, ou usando vários cabos em um único enlace de rede [14].
- Multiplexação por divisão de frequência (FDM) – particiona o espectro de frequência disponível em um conjunto de canais independentes. A utilização de FDM em uma rede ótica é denominada multiplexação por divisão de comprimento de onda (WDM) que permite uma fibra levar tráfego em muitos comprimentos de onda distintos [14].
- Multiplexação por divisão de tempo (TDM) – divide o domínio de tempo da banda em intervalos de duração fixa. Utilizando a TDM, múltiplos sinais podem compartilhar um determinado comprimento de onda se eles não se sobrepõem no tempo [14].

Embora a maior parte das pesquisas sobre *Traffic Grooming* na literatura se concentre na combinação eficiente de tráfegos de baixa velocidade em canais WDM de alta capacidade usando TDM como técnica de multiplexação, a idéia genérica de *grooming* pode ser aplicada a qualquer rede ótica através da utilização de uma das várias técnicas de multiplexação mencionadas acima.

A área de *Traffic Grooming* é formada por um rico conjunto de problemas, incluindo planejamento de redes, projeto de topologias (ambos baseados em tráfego estático) e provisionamento dinâmico de circuitos (baseado em tráfego dinâmico). O TGP baseado em tráfego estático é essencialmente um problema de otimização. Ele pode ser visto como um problema dual dependendo da perspectiva que se quer adotar. Uma primeira perspectiva é aquela em que, para um dado conjunto de requisições de tráfego, procura-se satisfazer todas as requisições de tráfego bem como minimizar o custo total da rede. Já o problema dual correspondente é aquele em que, para uma dada limitação de recursos e um dado conjunto de requisições de tráfego, procura-se maximizar o *throughput* da rede, ou melhor, a quantidade total de tráfego que é transportada com sucesso pela rede.

Nos últimos anos houve um crescente aumento da atividade de pesquisa sobre o TGP. Pesquisadores têm percebido que o TGP apresenta uma grande importância (teórica e prática) no planejamento e implementação de redes WDM. A seguir, o estado da arte desta área de pesquisa é revisado no tocante às redes em anel e redes irregulares.

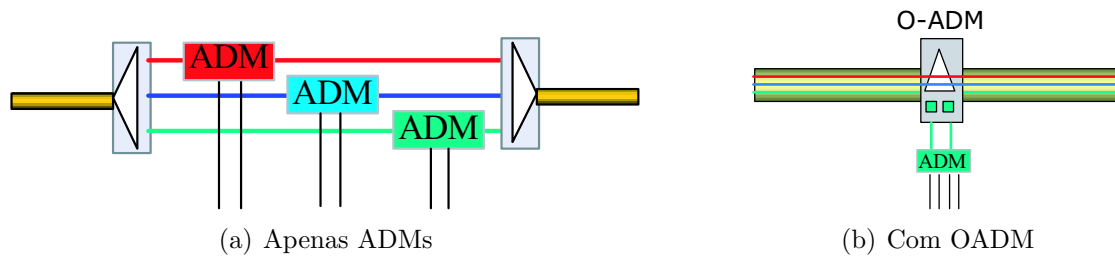


Figura 2.1: Arquiteturas de Elementos de Anel SONET+WDM (adaptada de [45])

2.3.2 TGP nas Redes em Anel

Redes óticas em anel utilizando tecnologia SONET/SDH representam a opção de infra-estrutura ótica mais amplamente usada no início do século XXI. Em uma rede em anel SONET/SDH, a WDM é utilizada principalmente como uma tecnologia de transmissão ponto a ponto. Cada comprimento de onda em tal rede SONET+WDM é operado a uma taxa de transmissão de alta capacidade como, por exemplo, OC-48 ou OC-192². Os esquemas de multiplexação por divisão do tempo (TDM) de um sistema SONET permitem que um canal de alta velocidade OC-N leve múltiplos canais OC-M (em que M é menor ou igual a N). A razão entre N e o valor menor de M transportado pela rede é chamada taxa de *grooming* (*grooming ratio*). Além disso, multiplexadores de inserção-retirada eletrônicos (*Add-Drop Multiplexers* – ADMs) são usados para inserir (retirar) tráfegos nos canais de alta velocidade pelos elementos da rede.

Em uma rede SONET tradicional (ver Figura 2.1(a)), um ADM é necessário para cada comprimento de onda em todo elemento da rede para que se realize a inserção/retirada de tráfego naquele comprimento de onda em particular. Com o progresso da tecnologia WDM já no início do século XXI mais de uma centena de comprimentos de onda podiam ser suportados simultaneamente por uma única fibra. Dessa forma, é muito dispendioso instalar a mesma quantidade de ADMs (cada um deles possui um custo significativo) em todos os elementos da rede, visto que grande parte do tráfego está apenas passando através de um elemento intermediário.

Através do uso de novos componentes óticos (ver Figura 2.1(b)) tais como multiplexadores óticos de inserção-retirada (*Optical Add-Drop Multiplexers* – OADM), também chamados de multiplexadores de inserção-retirada de comprimentos de onda (*Wavelength Add-Drop Multiplexers* – WADM), é possível que um elemento da rede insira e/ou retire apenas aqueles comprimentos de onda que transportam tráfego originado e/ou destinado a ele, fazendo com que a maior parte dos comprimentos de onda siga seu caminho sem interferência. Isto é realizado de uma forma “passiva” sem o envolvimento de componentes eletro-eletrônicos (ou melhor, através de elementos óticos). Neste tipo de rede, um *lightpath* representa um circuito virtual totalmente ótico (usando um único comprimento de onda) entre os componentes eletrônicos de um par de elementos da rede.

Comparado aos recursos necessários para o estabelecimento de um canal de comprimento de onda, o custo dos ADMs constitui a parcela dominante no custo total

²OC-48 equivale a 2,5 Gbps; enquanto OC-192 equivale a 10 Gbps.

de uma rede SONET+WDM. Conseqüentemente, uma seleção adequada dos locais de instalação dos ADMs, bem como um arranjo cuidadoso de quais comprimentos de onda serão inseridos/retirados ou não em cada um dos elementos da rede pode reduzir drasticamente o custo total da mesma (ver exemplo apresentado na subseção 1.1.3). Pode-se, então, dizer que os problemas a serem tratados são, para um dado conjunto de requisições de tráfego de baixa velocidade: estabelecer quais e de que forma tais tráfegos devem ser combinados durante seu transporte através da rede, determinar quais comprimentos de onda devem ser utilizados durante esse transporte, bem como quais comprimentos de onda devem ser inseridos/retirados em cada um dos elementos da rede além de, finalmente, se determinar quantos ADMs são necessários em cada um dos elementos.

TGP *Single-Hop* em Anéis

Considere que seja dada uma rede em anel SONET+WDM em que conexões de baixa velocidade OC-M devem ser combinadas para transporte através de canais de comprimento de onda de alta capacidade OC-N, onde $N \geq M$. Assuma, ainda, que não há conversor de comprimento de onda em qualquer elemento da rede, isto é, o tráfego em um comprimento de onda não pode ser comutado para outros comprimentos de onda. Com base neste modelo de rede, dada uma matriz de tráfego, satisfazer todas as requisições de tráfego bem como minimizar o número total de ADMs é um problema de projeto de rede e foi estudado extensivamente na literatura, sob a denominação de problema de *Traffic Grooming Single-Hop (TGP-SH)*.

Foi provado em [30, 154] que o problema de *Traffic Grooming* genérico é \mathcal{NP} -difícil. Os autores em [156] formularam este problema de otimização como um programa linear inteiro (PLI). Quando o tamanho da rede é pequeno, pacotes de otimização disponíveis comercialmente podem ser usados para resolver o PLI e obter uma ótima solução. A formulação em [156] pode ser aplicada tanto a situações envolvendo tráfego uniforme ou não, como também para redes em anel SONET+WDM unidirecionais e bidirecionais. A limitação da abordagem em [156] está ligada ao fato de que a quantidade de variáveis e restrições aumenta de forma explosiva com o aumento do tamanho das redes. Dessa forma, a complexidade computacional torna tal abordagem difícil de ser utilizada em redes com tamanhos mais realistas. Através da relaxação (remoção) de algumas restrições da formulação em [156], é possível se obter alguns resultados quase-ótimos para redes de tamanho razoável. Além disso, os resultados de [156] forneceram “insights” para o desenvolvimento de boas heurísticas para redes de grande porte.

Em [142, 154, 164] são apresentados limites inferiores para soluções de determinados modelos de tráfego (uniforme e não uniforme) e de rede (anel unidirecional e bidirecional). Tais limites inferiores podem ser usados para se avaliar o desempenho de algoritmos heurísticos para o problema. Na maioria das abordagens heurísticas, o TGP-SH é dividido em vários subproblemas que são resolvidos separadamente. Exemplos destas heurísticas podem ser encontrados em [30, 54, 142, 154, 156, 162, 164]. Tais algoritmos heurísticos lançam mão de abordagens gulosas, de meta-heurísticas (*simulated annealing*) e de algoritmos genéticos, entre outros.

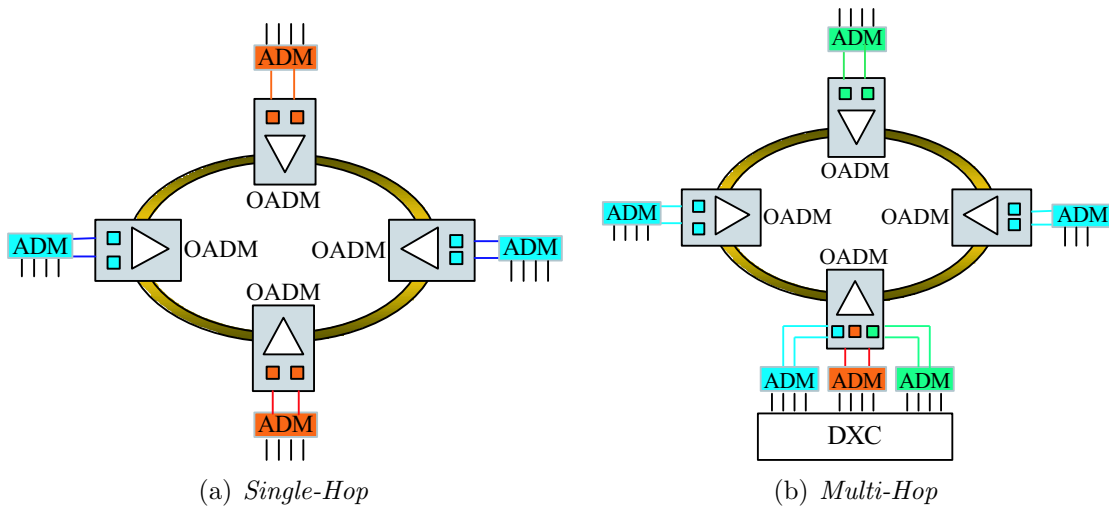


Figura 2.2: Arquitetura *Single-Hop* \times *Multi-Hop* (adaptada de [45])

TGP *Multi-Hop* em Anéis

No **TGP-SH**, o tráfego não pode ser comutado entre diferentes comprimentos de onda. Cada elemento da rede utiliza um OADM para inserir e/ou retirar comprimentos de onda específicos (ver Figura 2.2(a)). Deve haver, então, o estabelecimento de um *lightpath* entre cada par de elementos para os quais exista uma requisição de tráfego. Uma outra alternativa para a arquitetura de rede foi proposta em [55, 142], na qual há alguns elementos da rede equipados com comutadores (*Digital Crossconnects* – DXCs). A Figura 2.2(b) exibe um exemplo desta arquitetura. Este tipo de elemento é geralmente denominado de *hub*. No *hub*, um tráfego transportado através de um dado comprimento de onda (ou *time-slot*) pode ser comutado para qualquer outro comprimento de onda (ou *time-slot*). Uma vez que o tráfego precisa ser convertido do domínio óptico para o eletrônico no *hub* quando a troca de comprimento de onda (*time-slot*) ocorre (o que corresponde ao estabelecimento de pelo menos dois *lightpaths* distintos), esta abordagem dá origem a um problema denominado problema de *Traffic Grooming Multi-Hop* (**TGP-MH**). Dependendo da implementação, pode haver um único *hub* na rede ou vários deles. Um caso especial é aquele em que todo elemento da rede é um *hub*, isto é, há um DXC em todo elemento da rede. Este tipo de rede é denominada rede em anel WDM ponto-a-ponto (*Point-to-Point WDM* – PPWDM) [55].

O trabalho em [55] fornece uma análise teórica e compara os custos de uma rede em anel PPWDM, com um anel SONET+WDM sem *hub* e com um anel SONET+WDM com um ou vários *hubs*. Os autores de [156], por sua vez, compararam o desempenho da rede obtido na solução **TGP-SH** com aquele obtido pela resolução do **TGP-MH** (com um único *hub*) através de simulação. Os resultados indicam que, quando a taxa de *grooming* é alta, a abordagem *Multi-Hop* tende a usar um número menor de ADMs, mas quando essa taxa é pequena, a abordagem *Single-Hop* tende a ser melhor (usar menos ADMs) sendo que, em geral, a abordagem *Multi-Hop* utiliza um número maior de comprimentos de onda que a abordagem *Single-Hop* [156].

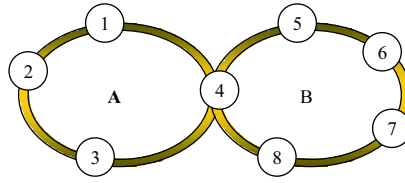


Figura 2.3: Exemplo de Anéis Interconectados (adaptada de [45])

TGP Dinâmico em Anéis

Em vez de se usar uma única matriz de tráfego estática para caracterizar as requisições de tráfego, também é possível descrevê-las através de um conjunto de matrizes de tráfego. O padrão de tráfego pode mudar dentro deste conjunto de matrizes durante um certo período de tempo, por exemplo, ao longo de um dia ou um mês. A rede precisa ser reconfigurada quando o padrão de tráfego se altera de uma matriz para outra matriz do conjunto. O problema de projeto de rede para suportar qualquer matriz de tráfego do conjunto de matrizes (de forma que não haja bloqueio de conexões) bem como para minimizar o custo total da rede é conhecido como **TGP** dinâmico e em [18] ele foi estudado para redes em anel SONET+WDM.

Diferentemente do problema de provisionamento dinâmico para uma rede irregular WDM que será abordado mais adiante, o **TGP** dinâmico proposto em [18] é mais próximo de um problema de projeto de rede em que se leva em consideração questões sobre a reconfiguração da mesma³. Os autores de [18] formularam o **TGP** dinâmico em uma rede em anel SONET+WDM como um problema de emparelhamento em um grafo bipartite e forneceram vários métodos para redução do número de ADMs, além das condições necessárias e suficientes para que uma rede possa suportar um determinado conjunto de padrões de tráfego (conjunto de matrizes). Por fim, em [163] uma abordagem para o **TGP** dinâmico através de um algoritmo genético para anéis SONET+WDM unidirecionais é apresentada e analisada.

TGP em Anéis Interconectados

A maioria das pesquisas sobre **TGP** para redes em anel SONET+WDM assumem que a topologia da rede consiste em um único anel. Os autores de [155] estenderam o problema para uma topologia de anéis interconectados (ver Figura 2.3). Atualmente as redes de transporte de alta velocidade (*backbone networks*) são formadas principalmente por uma série de anéis interconectados. A extensão dos estudos sobre **TGP** de topologias com um único anel para topologias com vários anéis interconectados é muito útil para que os operadores de rede possam projetar suas redes e gerenciar o tráfego das mesmas.

As diferentes arquiteturas possíveis para os elementos de interconexão (ver Figuras 2.4(a), 2.4(b) e 2.4(c)) entre os anéis irão adicionar diferentes restrições ao **TGP**. O trabalho em [155] apresenta uma formulação de programação linear inteira (PLI) para o **TGP** em uma topologia com anéis interconectados, além de propor um algoritmo heurístico para solucionar o problema para redes de tamanho real. Os resultados obtidos comparam várias estratégias de interconexão e taxas de *grooming*.

³Denominado, por vezes, como *Traffic Grooming and Reconfiguration Problem* – **TGRP**.

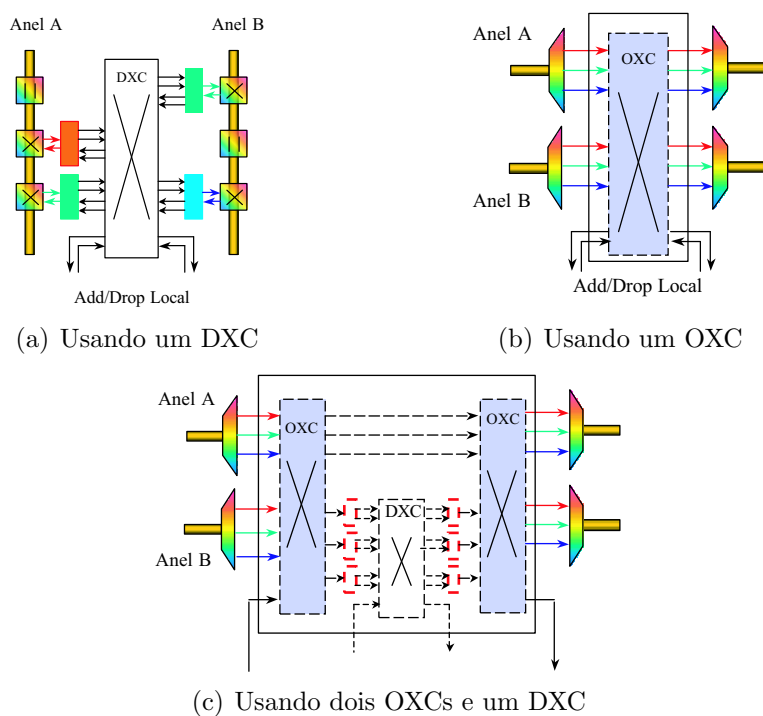


Figura 2.4: Arquiteturas dos Elementos de Interconexão (adaptada de [45])

Quando o número de anéis e o número de elementos de interconexão aumenta, a topologia da rede tende a se tornar uma topologia irregular, o que justifica o crescente interesse atual na resolução do **TGP** para redes com topologia irregular.

2.3.3 TGP nas Redes Irregulares

A maioria dos trabalhos sobre **TGP** encontrados na literatura enfocam redes cuja topologia é em anel. Recentemente, o **TGP** para redes irregulares WDM começou a atrair mais atenção dos pesquisadores. Nesta seção, alguns trabalhos recentes sobre este assunto são revisados.

Aprovisionamento/Reserva de Recursos de Rede

Embora as redes em anéis SONET/SDH (interconectados ou não) tenham sido usadas como infra-estrutura da primeira geração de redes óticas, elas possuem algumas limitações que dificultam sua escalabilidade e a acomodação do tráfego crescente da *Internet*. Espera-se que a próxima geração de redes óticas seja uma rede irregular WDM roteada “inteligentemente” através do uso de diferentes comprimentos de onda. Esta rede deverá permitir o provisionamento automático de largura de banda de forma rápida e conveniente (*point-and-click*) e mecanismos eficientes de proteção; além disso, ela estará baseada em uma topologia irregular, cuja expansão será muito mais fácil (resultando em uma melhoria da escalabilidade).

Quando tal rede for construída, o problema de se estabelecer a acomodação das requisições de tráfego eficazmente constituirá um problema de provisionamento (ou reserva) de recursos de rede. Uma requisição de tráfego pode ser estática (medida

através de uma ou várias matrizes de tráfego fixo) ou dinâmica (medida através de estatísticas sobre a taxa de chegada e o tempo de conexão das requisições). O trabalho em [169], baseado em requisições de tráfego estáticas, discute uma arquitetura para os elementos em uma rede irregular WDM com capacidade de realizar combinação de tráfegos de baixa velocidade (*grooming*).

Em tal arquitetura, um comutador ótico (*Optical Crossconnect* – OXC) com funcionalidades de comutação e multiplexação hierarquizadas é utilizado. Em vez de se usar um sistema de comutação de comprimentos de onda separado de um sistema de *grooming*, tal OXC pode suportar tráfegos de baixa velocidade diretamente e pode combiná-los em canais de comprimento de onda de alta capacidade através do uso de circuitos de *grooming* (*G-Fabric*) e de um conjunto de *transceivers* (transmissores e receptores óticos) embutidos no próprio equipamento (ver Figura 2.5). Esse tipo de OXC é denominado *Grooming OXC* (GOXC) ou comutador de comprimento de onda e *grooming* (*Wavelength-Grooming Crossconnect* – WGXC) [149]. Em uma rede equipada com um GOXC em cada um de seus elementos, o circuito de *grooming* e o tamanho do conjunto de *transceivers* fornecem um outro conjunto de restrições ao desempenho da rede, além das restrições associadas aos comprimentos de onda. Isto se assemelha às restrições sobre os ADMs para o TGP nas redes em anel SONET+WDM.

O conjunto de *transceivers* usado em um GOXC pode ser fixo ou ajustável (aqueles que podem receber/transmitir qualquer um dos comprimentos de onda utilizados na rede). Os autores em [169] consideraram um conjunto de matrizes de tráfego estáticas com as demandas de tráfego da rede. Cada matriz de tráfego do conjunto representa uma classe de requisições de baixa velocidade em particular. Para um dado conjunto de restrições sobre os recursos da rede e dadas as requisições de tráfego (conjunto de matrizes), o trabalho em [169] estuda como maximizar o *throughput* da rede. Como dito anteriormente, os problemas de se minimizar o custo e de se maximizar o *throughput* da rede representam duas perspectivas diferentes para o mesmo problema. Os autores em [169] formularam o problema como um PLI e utilizaram uma pequena rede para demonstrar os resultados obtidos pela formulação. Além disso, são propostos em [169], com base nas observações obtidas a partir desses resultados, dois algoritmos heurísticos para lidar com redes maiores e análises são realizadas para diferentes cenários, tais como: *Multi-Hop* \times *Single-Hop* e o uso de *transceivers* ajustáveis \times fixos.

Diferentemente do trabalho em [169], os trabalhos em [149, 148] consideram um padrão de tráfego dinâmico para uma rede irregular WDM. O trabalho em [149] propôs um esquema de controle de admissão de conexão (CAC) para assegurar que a rede irá tratar toda requisição de maneira justa (*fairly*). Foi observado em [149] que, quando a maioria dos elementos da rede tem capacidade de realizar *grooming*, os pedidos de conexão de alta velocidade terão probabilidade de bloqueio mais alta que os pedidos de conexão de baixa velocidade caso não haja um controle justo (equânime) na admissão das conexões. Um esquema de CAC é necessário para garantir que todas as classes de pedidos de conexão irão ter probabilidades de bloqueio semelhantes. O trabalho em [148] propôs um modelo teórico para calcular a probabilidade de bloqueio para redes irregulares WDM com capacidade limitada de *grooming*.

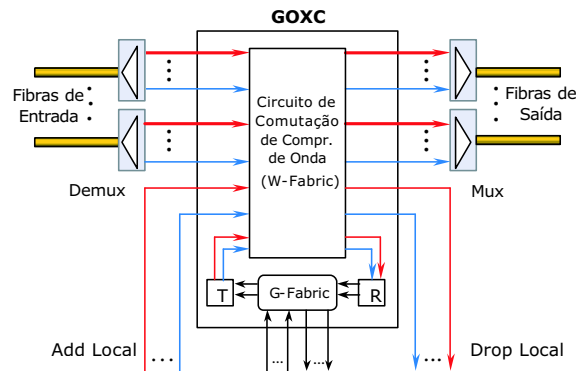


Figura 2.5: Arquitetura de OXC que realiza *Grooming* (adaptada de [45])

O trabalho em [169] assumiu que todo nó da rede é um GOXC e que a capacidade de *grooming* é limitada pelo circuito de *grooming* e pelo conjunto de *transceivers* em cada elemento da rede. O trabalho em [149] assumiu que apenas alguns dos elementos da rede são GOXCs e que não há nenhum limite sobre a capacidade de *grooming* dos mesmos.

Planejamento e Projeto de Rede

Diferentemente do problema de provisionamento de rede discutido acima, o trabalho em [36] estudou como planejar e projetar uma rede irregular WDM com base em uma previsão das requisições de tráfego, incluindo tanto o projeto da topologia física quanto o projeto da topologia virtual. O problema pode ser descrito da seguinte forma: dada uma previsão do tráfego (estática) e os elementos da rede (localizações), determinar como conectar tais elementos utilizando fibras óticas e OXCs e rotear o tráfego de modo a satisfazer todas as requisições de demanda bem como minimizar o custo total da rede. O custo total é mensurado a partir do custo da fibra, do custo associado ao OXC ou DXC (variando de acordo com o número de portas de entrada/saída utilizadas), além do custo de instalação do sistema WDM utilizado na rede.

Os autores em [36] abordam este problema de planejamento e projeto de rede formulando-o como um PLI. Dois algoritmos heurísticos foram propostos separadamente para o projeto de redes irregulares e para o projeto de rede em anel (em verdade, para projetar a rede com uma topologia irregular ou com uma topologia em anéis interconectados). Os autores compararam os resultados obtidos entre o projeto de uma rede irregular e o projeto de uma rede em anel e foram capazes de concluir que: (a) o projeto com uma topologia irregular apresenta um custo mais atrativo para os casos em que as distâncias entre os nós são razoavelmente grandes; (b) para as tecnologias de anel como OC-192 BLSR⁴, o uso de WDM só resulta em redução de custo quando as distâncias são suficientemente grandes; e (c) a função de custo adotada pode não refletir muito bem as distâncias para as tecnologias de anel [36].

⁴*Bidirectional Line-Switched Ring.*

TGP com Requisitos de Proteção

Já se demonstrou que para as redes em anel SONET+WDM é possível se adotar esquemas de proteção confiáveis não havendo, neste caso, nenhuma necessidade de se considerar os requisitos de proteção separadamente do TGP.

Em uma rede irregular WDM, vários esquemas de proteção podem ser usados dependendo da preferência do operador de rede e das exigências do cliente. Embora esquemas de proteção em redes irregulares WDM tenham sido estudados amplamente, o estudo de mecanismos de proteção em conjunto com o TGP representa uma área de pesquisa em aberto necessitando ser explorada adequadamente. Os tráfegos de baixa velocidade podem requerer diferentes larguras de banda bem como necessitar de diferentes serviços (esquemas) de proteção. Além disso, tais tráfegos podem utilizar de mecanismos de proteção tanto da camada eletrônica como da camada ótica da rede.

Dada uma matriz de tráfego estática e o tipo (esquema) de proteção para cada requisição (nenhuma proteção, 1+1, entre outros), os autores em [95] estudaram como satisfazer a demanda por largura de banda de tais conexões e atender as exigências de proteção procurando, ao mesmo tempo, minimizar o custo da rede. Nesse caso, o custo da rede é determinado pelo custo de transmissão e de comutação assemelhando-se ao descrito na seção anterior. A exigência de largura de banda de uma conexão pode ser uma fração de um canal de comprimento de onda e algumas conexões podem ser transportadas parcialmente pela camada eletrônica da rede. Os autores de [95] procuraram mostrar que haverá uma redução do custo da rede caso os tráfegos sejam combinados (*grooming*) para o transporte através da camada ótica ao invés de transportá-los pela camada eletrônica simplesmente. Uma formulação para o problema (PLI) é dada em [95] e uma heurística muito simples é proposta.

TGP com *Multicast*

Aplicações de *multicast* como vídeo sob demanda e jogos interativos estão se tornando cada vez mais populares. É razoável supor que num futuro próximo haverá mais aplicações de *multicast* como vídeo conferência, realidade virtual, entre outras.

Multicasting ótico utilizando *lighttrees* [139] pode ser uma boa solução para estes casos. Uma vez que cada comprimento de onda pode ter uma capacidade de até 10 Gbps (40 Gbps, no futuro), múltiplas sessões de *multicast* podem ser combinadas (*grooming*) de modo a compartilhar a capacidade do mesmo canal de comprimento de onda. Neste caso, os *lightpaths* (ou *lighttrees*) podem ser estabelecidos para acomodar as requisições que possuem uma necessidade de largura de banda menor que a capacidade de um comprimento de onda.

O trabalho em [143] relata um estudo preliminar sobre TGP com *multicast* em redes irregulares WDM. O problema é definido da seguinte forma: dado um conjunto de sessões de *multicast* com diferentes exigências de capacidade, satisfazer todas as sessões de *multicast* e, ao mesmo tempo, minimizar o custo da rede. Os autores de [143] apresentam uma formulação (PLI) para este problema e alguns resultados baseados em matrizes de tráfego e topologias de rede simples. Infelizmente a abordagem adotada em [143] (uso de pacotes comerciais para resolução do PLI proposto) se mostra de difícil aplicação para redes de tamanho real. Consequentemente, al-

goritmos mais simples e eficientes precisam ser desenvolvidos para se obter soluções ótimas (ou quase-ótimas). O **TGP** com *multicast* é uma área de pesquisa muito recente e que provavelmente irá receber mais atenção no futuro.

Extensões de Protocolos e Algoritmos para Controle

O **TGP** é um problema muito importante cuja solução irá permitir o desenvolvimento de uma rede de transporte ótica WDM “inteligente”. O plano de controle de tal rede está sendo unificado, e é conhecido como *Generalized Multi-Protocol Label Switching* (GMPLS) [161]. O propósito deste plano de controle é prover um esquema de provisionamento/sinalização de conexões ponto-a-ponto (circuito virtual) de forma inteligente e automática entre diferentes domínios de rede. Técnicas de multiplexação diferentes podem ser utilizadas (tais como TDM, WDM e SDM) para o estabelecimento de tais conexões, e bons esquemas de *grooming* são necessários para se alocar recursos de rede de maneira adequada.

Há três componentes do plano de controle que precisam ser projetados cuidadosamente para suportar o *Traffic Grooming*: protocolo de localização de recursos (*resource-discovery protocol*), protocolo de sinalização (*signaling protocol*) e algoritmos para cálculo de rotas (*path-computation algorithms*). Vários protocolos de localização de recursos com base em extensões de protocolos de estado de enlace (OSPF, IS-IS) [17, 22, 88] e protocolos de gerência de enlace [94] foram propostos pelo IETF. Extensões do protocolo MPLS, por sua vez, têm sido propostas como protocolos de sinalização. Um tópico em aberto é o projeto de algoritmos eficientes para cálculo de rotas. Os trabalhos em [167, 168] apresentam resultados preliminares sobre esquemas de provisionamento *online* para conexões com largura de banda de diferentes granularidades (**TGP** dinâmico) para redes irregulares WDM que utilizam GMPLS. Várias políticas de *grooming* são propostas nesses trabalhos. Os resultados desses estudos indicam que, para se alcançar um bom desempenho em um ambiente dinâmico, diferentes políticas de *grooming* e algoritmos para cálculo de rotas precisam ser usados dependendo dos diferentes estados em que a rede se encontra.

2.4 Formulações Existentes para o TGP

Nas subseções 2.3.2 e 2.3.3, optou-se por discutir as diversas abordagens para o **TGP** (e suas variações) encontradas na literatura para redes óticas em anel e irregulares, respectivamente; contudo, nenhuma descrição formal das formulações de programação matemática para o problema foi apresentada. Esta seção se incumbe de realizar tal exposição com o intuito de permitir uma melhor compreensão das similaridades e diferenças entre as formulações existentes na literatura e a “nova” formulação do **TGP** que será introduzida no próximo capítulo.

Primeiramente, cabe destacar a formulação de programação linear inteira apresentada em [156] que pode ser vista como uma síntese dos esforços, até então, para modelagem e resolução do **TGP** em redes com topologia em anel, constituindo, segundo seus autores, a primeira especificação formal (via programação matemática) do problema.

Na realidade em [156] são propostas duas formulações distintas (uma para o **TGP-SH** e outra para o **TGP-MH** com um único *hub*) que são apresentadas a seguir. Antes, porém, deve-se introduzir a seguinte notação utilizada em tais formulações de programação matemática:

N	representa o número de nós da rede;
W	representa o número de comprimentos de onda disponíveis na rede;
C	representa a taxa de <i>grooming</i> (<i>grooming ratio</i> - ver início da subseção 2.3.2);
T	representa uma matriz de tráfego não uniforme, em que t_{ij} é a quantidade de tráfego que sai do elemento de rede i com destino a j ;
${}^dV_{ij}^{cw}$	representa uma conexão virtual do nó i para o nó j utilizando o círculo ⁵ c no comprimento de onda w , em que d indica o sentido da conexão (horário ou anti-horário);
ADM_i^w	representa o número de ADMs no nó i para o comprimento de onda w ;
e	representa um arco qualquer da rede física;
${}^de^{cw}$	representa um arco no sentido d no comprimento de onda w usando o círculo (ou subcanal) c (${}^de^{cw} \in {}^dV_{ij}^{cw}$, caso tal arco seja utilizado pela conexão virtual ${}^dV_{ij}^{cw}$).

Sendo assim, a formulação de programação linear inteira do **TGP-SH** para um anel bidirecional apresentada em [156] é dada por:

$$(TGPSH_A) \quad \min \sum_i \sum_w ADM_i^w \quad (2.1a)$$

sujeito a:

$$\sum_w \sum_c \sum_d {}^dV_{ij}^{cw} = t_{ij} \quad \forall i, j \quad (2.1b)$$

$$\sum_{{}^de^{cw} \in {}^dV_{ij}^{cw}} {}^dV_{ij}^{cw} \leq 1 \quad \forall d, e, c, w \quad (2.1c)$$

$$\sum_c \sum_j {}^dV_{ij}^{cw} \leq C \cdot ADM_i^w \quad \forall d, i, w \quad (2.1d)$$

$$\sum_c \sum_i {}^dV_{ij}^{cw} \leq C \cdot ADM_j^w \quad \forall d, j, w \quad (2.1e)$$

$${}^dV_{ij}^{cw} \in \{0, 1\} \quad \forall d, i, j, c, w \quad (2.1f)$$

$$ADM_i^w \in \{0, 1\} \quad \forall i, w \quad (2.1g)$$

A função objetivo (2.1a) contabiliza o número total de ADMs utilizados. As restrições (2.1b) garantem que o número total de conexões entre cada par de nós em todos é igual a demanda especificada na matriz de tráfego. A capacidade de cada círculo (ou subcanal) é limitada ao transporte de uma unidade de tráfego pelas restrições (2.1c). As restrições (2.1d) e (2.1e) especificam que o número de conexões iniciadas e terminadas, respectivamente, em um nó do anel é limitado pela capacidade de comutação eletrônica (número de ADMs) existentes em tal elemento.

⁵Um círculo representa um anel virtual (ou subcanal) estabelecido em um *time-slot* de um comprimento de onda e que possui capacidade para transportar uma unidade de tráfego.

Finalmente, as restrições (2.1f) e (2.1g) definem as variáveis da formulação como binárias.

Além da formulação do **TGP-SH** para um anel bidirecional, os autores em [156] apresentam uma formulação do **TGP-MH** com um único *hub* para um anel unidirecional. Para tanto, torna-se necessária a introdução da seguinte notação adicional:

- O_i^{cw} representa a conexão virtual que se inicia no nó i e termina no *hub* utilizando o círculo (subcanal) c na comprimento de onda w ;
- I_i^{cw} representa a conexão virtual que se inicia no *hub* e termina no nó i utilizando o círculo (subcanal) c na comprimento de onda w ;
- V_{ij}^{cw} representa uma conexão virtual do nó i para o nó j utilizando o círculo c no comprimento de onda w ;
- e^{cw} representa um arco no comprimento de onda w usando o círculo c ($e^{cw} \in V_{ij}^{cw}$, caso ele seja utilizado pela conexão virtual V_{ij}^{cw}).

Dessa forma, a formulação de programação linear inteira do **TGP-MH** com um único *hub* para um anel unidirecional apresentada em [156] é dada da seguinte forma:

$$(TGPMH_A) \quad \min \sum_i \sum_w ADM_i^w \quad (2.2a)$$

sujeito a:

$$\sum_w \sum_c \left(\sum_{j|j>i} V_{ij}^{cw} + O_i^{cw} \right) = \sum_j t_{ij} \quad \forall i \quad (2.2b)$$

$$\sum_w \sum_c \left(\sum_{i|j>i} V_{ij}^{cw} + I_j^{cw} \right) = \sum_i t_{ij} \quad \forall j \quad (2.2c)$$

$$\sum_{e^{cw} \in V_{ij}^{cw}} V_{ij}^{cw} + \sum_{i|i < e^{cw}} O_i^{cw} + \sum_{j|j > e^{cw}} I_j^{cw} \leq 1 \quad \forall e, c, w \quad (2.2d)$$

$$\sum_c \sum_j V_{ij}^{cw} + \sum_c O_i^{cw} \leq C \cdot ADM_i^w \quad \forall i, w \quad (2.2e)$$

$$\sum_c \sum_j V_{ji}^{cw} + \sum_c I_i^{cw} \leq C \cdot ADM_i^w \quad \forall i, w \quad (2.2f)$$

$$V_{ij}^{cw} \in \{0, 1\} \quad \forall i, j, c, w \quad (2.2g)$$

$$O_i^{cw} \in \{0, 1\} \quad \forall i, c, w \quad (2.2h)$$

$$I_i^{cw} \in \{0, 1\} \quad \forall i, c, w \quad (2.2i)$$

$$ADM_i^w \in \{0, 1\} \quad \forall i, w \quad (2.2j)$$

A função objetivo (2.2a) contabiliza o número total de ADMs utilizados. As restrições (2.2b) e (2.2c) garantem que o número total de conexões entre cada par de nós em todos é igual a demanda especificada na matriz de tráfego. Vale dizer que a condição $j > i$ será satisfeita por aquelas conexões virtuais que se iniciam em i e terminam em j sem passar pelo *hub*. A capacidade de cada círculo (ou subcanal) é limitada ao transporte de uma unidade de tráfego pelas restrições (2.2d), em que a condição $i < e^{cw}$ será atendida por todo nó situado entre o *hub* e o nó de início do arco e^{cw} (incluindo esse último). Similarmente, a condição $j > e^{cw}$ será atendida por todo nó entre o nó de destino do arco e^{cw} (inclusive) e o *hub*. As restrições (2.2e) e (2.2f) especificam que o número de conexões iniciadas e terminadas, respectivamente, em um nó do anel é limitado pela capacidade de comutação eletrônica (número de

ADMs) existentes em tal elemento. Finalmente, as restrições (2.2g), (2.2h), (2.2i) e (2.2j) definem as variáveis da formulação como binárias.

Conforme mencionado anteriormente, o **TGP** em redes irregulares tem recebido maior atenção recentemente. Em [169], os autores investigam o problema procurando maximizar o *throughput* da rede, ou melhor, a quantidade total de tráfego que é transportada com sucesso pela rede.

Em [169], duas formulações para redes com topologia irregular (uma para o **TGP-SH** e outra para o **TGP-MH**) são propostas. Antes, porém, deve-se introduzir a seguinte notação utilizada em tais formulações de programação matemática:

- N representa o número de nós da rede;
- W representa o número de comprimentos de onda;
- C representa a capacidade de cada comprimento de onda (*grooming ratio*);
- P_{mn} representa o número de fibras conectando o nó m ao nó n , $P_{nm} = 0$ para todo par (m, n) que não é fisicamente adjacente e $P_{mn} = P_{nm} = 1$ se e somente se existe uma conexão física direta entre os nós m e n ;
- P_{mn}^w representa o comprimento de onda w na fibra entre os nós m e n , em [169], $P_{mn}^w = P_{mn}$;
- TR_i representa o número de transmissores no nó i ;
- RR_i representa o número de receptores no nó i ;
- y representa a “granularidade” das requisições de baixa velocidade, os autores em [169] adotaram $y \in \{1, 3, 12, 48\}$, o que significa que as demandas de tráfego entre os nós podem ser requisições por conexões OC-1, OC-3, OC-12 e OC-48;
- Λ representa um conjunto de matrizes de tráfegos, $\Lambda = \{\Lambda_y\}$, em que y corresponde a qualquer um dos tráfegos de baixa velocidade, em [169], $y \in \{1, 3, 12, 48\}$;
- $\Lambda_{y,sd}$ representa o número de requisições por conexões OC- y entre o par de nós (s, d) ;
- t representa o índice de uma requisição de tráfego OC- y para um dado par de nós (s, d) , isto é, $t \in [1, 2, \dots, \Lambda_{y,sd}]$, por exemplo caso existam 10 requisições de tráfego OC-1 entre um par (s, d) então $t \in [1, 2, \dots, 10]$.

Além disso, as seguintes variáveis serão utilizadas nas formulações de programação matemática:

- V_{ij} representa o número de *lightpaths* entre o par de nós (i, j) ;
- V_{ij}^w representa o número de *lightpaths* entre o par de nós (i, j) utilizando o comprimento de onda w ;
- $P_{mn}^{ij,w}$ indica se um *lightpath* entre o par de nós (i, j) utilizando o comprimento de onda w passa pelo arco da rede física (m, n) ;
- $\lambda_{ij,y}^{sd,t}$ indica se o *lightpath* entre o par de nós (i, j) é utilizado no roteamento da t -ésima requisição de tráfego de baixa velocidade OC- y entre o par de nós (s, d) ;
- $S_{sd}^{y,t}$ indica se houve sucesso no roteamento da t -ésima requisição de tráfego de baixa velocidade OC- y entre o par de nós (s, d) .

Dessa forma, a formulação de programação linear inteira do **TGP-MH** para uma rede irregular apresentada em [169] é dada da seguinte forma:

$$(TGPMH_1) \quad \max \sum_{y,s,d,t} y \cdot S_{sd}^{y,t} \quad (2.3a)$$

sujeito a:

$$\sum_j V_{ij} \leq TR_i \quad \forall i \quad (2.3b)$$

$$\sum_i V_{ij} \leq RR_j \quad \forall j \quad (2.3c)$$

$$\sum_w V_{ij}^w = V_{ij} \quad \forall i, j \quad (2.3d)$$

$$\sum_m P_{mk}^{ij,w} - \sum_n P_{kn}^{ij,w} = 0 \quad \text{se } k \neq i \neq j, \quad \forall i, j, w, k \quad (2.3e)$$

$$\sum_m P_{mi}^{ij,w} = 0 \quad \forall i, j, w \quad (2.3f)$$

$$\sum_n P_{jn}^{ij,w} = 0 \quad \forall i, j, w \quad (2.3g)$$

$$\sum_n P_{in}^{ij,w} = V_{ij}^w \quad \forall i, j, w \quad (2.3h)$$

$$\sum_m P_{mj}^{ij,w} = V_{ij}^w \quad \forall i, j, w \quad (2.3i)$$

$$\sum_{i,j} P_{mn}^{ij,w} \leq P_{mn}^w \quad \forall m, n, w \quad (2.3j)$$

$$\sum_i \lambda_{ik,y}^{sd,t} - \sum_j \lambda_{kj,y}^{sd,t} = 0 \quad \text{se } k \neq s \neq d, \quad \forall s, d, k, y, t \quad (2.3k)$$

$$\sum_i \lambda_{is,y}^{sd,t} = 0 \quad \forall s, d, y, t \quad (2.3l)$$

$$\sum_j \lambda_{dj,y}^{sd,t} = 0 \quad \forall s, d, y, t \quad (2.3m)$$

$$\sum_i \lambda_{id,y}^{sd,t} = S_{sd}^{y,t} \quad \forall s, d, y, t \quad (2.3n)$$

$$\sum_j \lambda_{sj,y}^{sd,t} = S_{sd}^{y,t} \quad \forall s, d, y, t \quad (2.3o)$$

$$\sum_{y,s,d,t} y \cdot \lambda_{ij,y}^{sd,t} \leq C \cdot V_{ij} \quad \forall i, j \quad (2.3p)$$

$$V_{ij} \in \mathbb{Z}_+ \quad \forall i, j \quad (2.3q)$$

$$V_{ij}^w \in \mathbb{Z}_+ \quad \forall i, j, w \quad (2.3r)$$

$$P_{mn}^{ij,w} \in \{0, 1\} \quad \forall i, j, m, n, w \quad (2.3s)$$

$$\lambda_{ij,y}^{sd,t} \in \{0, 1\} \quad \forall i, j, s, d, y, t \quad (2.3t)$$

$$S_{sd}^{y,t} \in \{0, 1\} \quad \forall s, d, y, t \quad (2.3u)$$

A função objetivo (2.3a) contabiliza o volume total tráfego transportado através da rede. As restrições (2.3b) e (2.3c) asseguram que o número de *lightpaths*

entre o par de nós (i, j) é menor ou igual ao número de transmissores do nó i e de receptores do nó j . As restrições (2.3d) estabelecem que os *lightpaths* entre um par (i, j) podem ser formados por *lightpaths* em diferentes comprimentos de onda. As restrições (2.3e), (2.3f), (2.3g), (2.3h) e (2.3i) são restrições de conservação de fluxo e são responsáveis pelo roteamento dos *lightpaths* garantindo a *continuidade* de comprimento de onda. Já as restrições (2.3j) garantem que o comprimento de onda w em uma fibra entre o par de nós (m, n) só seja utilizado por no máximo um *lightpath* da topologia virtual. Analogamente, as restrições (2.3k), (2.3l), (2.3m), (2.3n) e (2.3o) também são restrições de conservação de fluxo; contudo, elas são responsáveis pelo roteamento das requisições de baixa velocidade através da topologia virtual. As restrições (2.3p) garantem que o tráfego agregado total transportado pelos *lightpaths* não excede a capacidade total disponível. Finalmente, as restrições (2.3q), (2.3r), (2.3s), (2.3t) e (2.3u) definem as variáveis da formulação.

Já a formulação de programação linear inteira do **TGP-SH** para uma rede irregular apresentada em [169] é dada por:

$$(TGPSH_1) \quad \max \sum_{y,s,d,t} y \cdot S_{sd}^{y,t} \quad (2.4a)$$

sujeito a:

$$\sum_j V_{ij} \leq TR_i \quad \forall i \quad (2.4b)$$

$$\sum_i V_{ij} \leq RR_j \quad \forall j \quad (2.4c)$$

$$\sum_w V_{ij}^w = V_{ij} \quad \forall i, j \quad (2.4d)$$

$$\sum_m P_{mk}^{ij,w} - \sum_n P_{kn}^{ij,w} = 0 \quad \text{se } k \neq i, j, \quad \forall i, j, w, k \quad (2.4e)$$

$$\sum_m P_{mi}^{ij,w} = 0 \quad \forall i, j, w \quad (2.4f)$$

$$\sum_n P_{jn}^{ij,w} = 0 \quad \forall i, j, w \quad (2.4g)$$

$$\sum_n P_{in}^{ij,w} = V_{ij}^w \quad \forall i, j, w \quad (2.4h)$$

$$\sum_m P_{mj}^{ij,w} = V_{ij}^w \quad \forall i, j, w \quad (2.4i)$$

$$\sum_{i,j} P_{mn}^{ij,w} \leq P_{mn}^w \quad \forall m, n, w \quad (2.4j)$$

$$\sum_{y,s,d,t} y \cdot S_{sd}^{y,t} \leq C \cdot V_{sd} \quad \forall s, d \quad (2.4k)$$

$$V_{ij} \in \mathbb{Z}_+ \quad \forall i, j \quad (2.4l)$$

$$V_{ij}^w \in \mathbb{Z}_+ \quad \forall i, j, w \quad (2.4m)$$

$$P_{mn}^{ij,w} \in \{0, 1\} \quad \forall i, j, m, n, w \quad (2.4n)$$

$$S_{sd}^{y,t} \in \{0, 1\} \quad \forall s, d, y, t \quad (2.4o)$$

Essa formulação ($TGPSH_M$) é muito similar a anterior ($TGPMH_M$) exceto pela ausência de roteamento das requisições de baixa velocidade através da topologia virtual (assume-se que uma conexão só pode ser transportada através de um único *lightpath*). Portanto, as restrições (2.3k)–(2.3p) foram substituídas pelas restrições (2.4k) e as restantes permanecem inalteradas.

Tanto as formulações propostas em [169] para redes irregulares como aquelas apresentadas em [156] para redes em anel só podem ser diretamente aplicadas para instâncias muito pequenas. De fato, tais formulações foram utilizadas por seus autores em conjunto com o pacote comercial *CPLEX* [76] e apenas instâncias de no máximo 6 nós puderam ser resolvidas com um esforço computacional aceitável. Dessa forma, inúmeras heurísticas (conforme já descrito anteriormente) passaram a ser desenvolvidas e testadas para o **TGP**; contudo, nenhum esforço para obtenção da solução ótima do problema foi realizado exceto pelo trabalho apresentado em [42, 44].

Este trabalho apresenta uma formulação de programação linear inteira para o **TGP** em um anel unidirecional a partir da qual se desenvolve um método para o cálculo de limites inferiores para o problema. Tais limites são obtidos através da decomposição do anel em subconjuntos conexos de nós para os quais a topologia virtual “localmente ótima” é adotada. Além disso, limites superiores são gerados através da aplicação de uma série de heurísticas. De modo a se obter uma solução ótima (ou pelo menos melhor que a disponível em um dado momento), um procedimento de atualização dos limites foi desenvolvido e avaliado [42, 44]. Entretanto, devido a especificidade dessa abordagem sua extensão para outras topologias de rede física não parece ser trivial.

De todo modo, em [45] o modelo apresentado em [42, 44] é revisto e estendido para redes irregulares sem que nenhum procedimento de resolução seja apresentado. Mesmo assim, optou-se por apresentar aqui apenas tal formulação mais geral, uma vez que a formulação em [42, 44] pode ser vista como uma “versão simplificada” da formulação de [45]. A notação utilizada em [45] é a seguinte:

N	representa o número de nós da rede;
W	representa o número de comprimentos de onda;
C	representa a capacidade de cada comprimento de onda (<i>grooming ratio</i>);
T	representa a matriz de tráfego, $T = \{t^{(sd)}\}$, em que $t^{(sd)}$ representa a quantidade de tráfego que sai do nó s com destino ao nó d ;
t_{ij}	representa a quantidade de tráfego agregado que é transportada por <i>lightpaths</i> entre o par de nós (i, j) , em que $t_{ij}^{(sd)}$ representa a porção deste tráfego relacionada a demanda $t^{(sd)}$;
b_{ij}	representa o número de <i>lightpaths</i> entre o par de nós (i, j) ;
$b_{ij}(l, m)$	representa o número de <i>lightpaths</i> entre o par (i, j) que atravessam conexão física (fibra) entre os nós l e m ;
$c_{ij}^{(k)}(l, m)$	indica que o <i>lightpath</i> entre o par de nós (i, j) utiliza o comprimento de onda k na conexão física (fibra) entre os nós l e m ;
p_{lm}	indica a existência ou não da conexão física (fibra) entre os nós l e m .

Sendo assim, a formulação de programação linear inteira do **TGP-MH** para uma rede irregular apresentada em [45] é dada por:

$$(\overline{TGPMH}_1) \quad \min \sum_{i,j} b_{ij} \quad (2.5a)$$

sujeito a:

$$b_{ij}(l, m) \leq b_{ij} \cdot p_{lm} \quad \forall i, j, l, m \quad (2.5b)$$

$$c_{ij}^{(k)}(l, m) \leq p_{lm} \quad \forall i, j, k, l, m \quad (2.5c)$$

$$\sum_{l=0}^{N-1} b_{ij}(i, l) - \sum_{l=0}^{N-1} b_{ij}(l, i) = b_{ij} \quad \forall i, j \quad (2.5d)$$

$$\sum_{l=0}^{N-1} b_{ij}(j, l) - \sum_{l=0}^{N-1} b_{ij}(l, j) = -b_{ij} \quad \forall i, j \quad (2.5e)$$

$$\sum_{l=0}^{N-1} b_{ij}(m, l) - \sum_{l=0}^{N-1} b_{ij}(l, m) = 0 \quad \text{se } m \neq i \neq j, \forall i, j, m \quad (2.5f)$$

$$\sum_{i,j} b_{ij}(l, m) \leq W \quad \forall l, m \quad (2.5g)$$

$$\sum_{k=0}^{W-1} c_{ij}^{(k)}(l, m) = b_{ij}(l, m) \quad \forall i, j, l, m \quad (2.5h)$$

$$\sum_{i,j} c_{ij}^{(k)}(l, m) \leq 1 \quad \forall k, l, m \quad (2.5i)$$

$$\sum_{l=0}^{N-1} c_{ij}(i, l) - \sum_{l=0}^{N-1} c_{ij}(l, i) \leq b_{ij} \quad \forall i, j, k \quad (2.5j)$$

$$\sum_{l=0}^{N-1} c_{ij}(j, l) - \sum_{l=0}^{N-1} c_{ij}(l, j) \geq -b_{ij} \quad \forall i, j, k \quad (2.5k)$$

$$\sum_{l=0}^{N-1} c_{ij}(m, l) - \sum_{l=0}^{N-1} c_{ij}(l, m) = 0 \quad \text{se } m \neq i \neq j, \forall i, j, k, m \quad (2.5l)$$

$$\sum_{j=0}^{N-1} t_{sj}^{(sd)} - \sum_{j=0}^{N-1} t_{js}^{(sd)} = t^{(sd)} \quad \forall s, d \quad (2.5m)$$

$$\sum_{j=0}^{N-1} t_{dj}^{(sd)} - \sum_{j=0}^{N-1} t_{jd}^{(sd)} = -t^{(sd)} \quad \forall s, d \quad (2.5n)$$

$$\sum_{j=0}^{N-1} t_{ij}^{(sd)} - \sum_{j=0}^{N-1} t_{ji}^{(sd)} = 0 \quad \text{se } i \neq s \neq d, \forall s, d, i \quad (2.5o)$$

$$\sum_{s,d} t_{ij}^{(sd)} = t_{ij} \quad \forall i, j \quad (2.5p)$$

$$t_{ij} \leq C \cdot b_{ij} \quad \forall i, j \quad (2.5q)$$

$$t_{ij} \in \mathbb{Z}_+ \quad \forall i, j \quad (2.5r)$$

$$b_{ij} \in \mathbb{Z}_+ \quad \forall i, j \quad (2.5s)$$

$$b_{ij}(l, m) \in \mathbb{Z}_+ \quad \forall i, j, l, m \quad (2.5t)$$

$$c_{ij}^{(k)}(l, m) \in \{0, 1\} \quad \forall i, j, k, l, m \quad (2.5u)$$

A função objetivo (2.5a) contabiliza o número total de *lightpaths*. As restrições (2.5b) e (2.5c) limitam a alocação de arcos da topologia virtual (isto é, pertencentes a algum *lightpath*) apenas àqueles existentes na topologia física. As restrições (2.5d), (2.5e) e (2.5f) são restrições de conservação de fluxo e são responsáveis pelo roteamento dos *lightpaths*. As restrições (2.5g) garantem que o número de *lightpaths* utilizando um mesmo arco da topologia física seja no máximo igual ao número de comprimentos de onda disponíveis. As restrições (2.5h), (2.5i), (2.5j), (2.5k) e (2.5l) garantem a resolução correta do subproblema de atribuição de comprimentos de onda. As restrições (2.5h) estabelecem que o número de comprimentos de onda alocados em um dado arco da topologia física deve ser igual ao número de *lightpaths* que se utilizam desse mesmo arco. Já as restrições (2.5i) evitam que o mesmo comprimento de onda seja utilizado por mais de um *lightpath* passando pelo mesmo arco da rede física (isto é, impedem o *waveclash*). Finalmente, as restrições (2.5j), (2.5k) e (2.5l) garantem a *continuidade* de comprimento de onda.

O roteamento de tráfegos através da topologia virtual é garantido pelas restrições (2.5m), (2.5n) e (2.5o), enquanto as restrições (2.5p) e (2.5q) estabelecem a ligação entre o subproblema de roteamento de tráfego e o subproblema de determinação e roteamento de *lightpaths*. Por fim, as restrições (2.5r), (2.5s), (2.5t) e (2.5u) definem as variáveis da formulação.

Por fim, cabe reafirmar o objetivo desta seção em realizar uma exposição das formulações existentes na literatura com o intuito de permitir uma melhor compreensão das similaridades e diferenças entre tais formulações e a “nova” formulação do **TGP** apresentada a seguir. Posto isto, o capítulo seguinte introduz uma representação em camadas da rede a partir da qual se desenvolve um conjunto de formulações de programação matemática para o **TGP** (e algumas de suas variações) que são utilizadas no restante deste trabalho.

Capítulo 3

Modelos Matemáticos

DE modo a utilizar mais eficazmente a largura de banda disponível, as recentes arquiteturas de redes óticas têm permitido que vários tráfegos independentes compartilhem a largura de banda disponibilizada por um *lightpath*. Caso a multiplexação (combinação) dos tráfegos de baixa velocidade seja realizada apenas na fronteira da rede de OXCs e o tráfego agregado seja roteado apenas no domínio ótico (de forma transparente), este problema é totalmente equivalente ao problema de *RWA*. Entretanto, geralmente não é possível se estabelecer um *lightpath* distinto entre cada par de elementos da rede, em virtude da existência de restrições sobre o número de comprimentos de onda disponíveis ou sobre o número de *transceivers* óticos existentes em cada elemento da rede (ver exemplo apresentado na subseção 1.1.3). Sendo assim, torna-se natural considerar que as redes óticas são constituídas por elementos capazes de realizar algum tipo de roteamento eletrônico, de modo que um *lightpath* possa ser retirado da rede ótica e todo o seu conteúdo (tráfegos de baixa velocidade que foram multiplexados previamente) ou, até mesmo parte dele, possa ser comutado/roteado através de um outro *lightpath* em direção ao seu destino.

A introdução de alguma capacidade de roteamento eletrônico nos elementos da rede traz consigo vantagens: primeiramente, ela é responsável por um aumento do grau de conectividade (na topologia virtual) dos elementos da rede, ao mesmo tempo em que ela pode significar uma redução drástica na quantidade de comprimentos de onda necessária para se transportar um dado conjunto de requisições de tráfego. Contudo, a introdução de componentes ativos (isto é, que realizam processamento eletrônico) e caros, tais como comutadores/roteadores eletrônicos de alto desempenho e *transceivers* óticos, conduz a um aumento do custo da rede.

Segundo os autores de [45], estas observações motivam a definição e estudo do **TGP** e devido à sua relação com o problema de projeto de topologias virtuais, o **TGP** poderia ser subdividido em três subproblemas distintos: determinação da topologia (virtual), ou melhor, cálculo dos *lightpaths* entre os elementos da rede; roteamento e atribuição de comprimentos de onda aos *lightpaths* (problema de *RWA*) e determinação do roteamento do tráfego através do conjunto de *lightpaths*. O objetivo do **TGP**, então, seria solucionar tais subproblemas ao mesmo tempo em que se minimiza uma função de custo definida sobre o conjunto de *lightpaths*.

Ainda segundo [45] e em consonância com diversas outras abordagens, três funções de custo são relevantes neste cenário e merecem ser discutidas:

1. Número total de *lightpaths*:

Neste modelo, o custo a ser minimizado representa a cardinalidade do conjunto de *lightpaths*. Esse modelo de custo é motivado pela observação de que cada *lightpath* necessita de um *transceiver* em cada uma de suas extremidades, além de requerer suficiente capacidade para rotear os tráfegos que o compõem. Desse modo, ao contabilizar o número de *lightpaths*, esta função de custo é capaz de capturar tanto o custo relativo aos LTEs quanto o custo associado a comutação/roteamento eletrônico para o atendimento de um conjunto de requisições de tráfego. Uma métrica alternativa seria o número total de LTEs, isto é, o número de ADMs para uma rede com topologia em anel ou o número de *transceivers* nos OXCs para uma rede com topologia arbitrária.

2. Quantidade total de tráfego roteado eletronicamente:

Em lugar de se contabilizar o número de *lightpaths* manipulados, este modelo avalia a quantidade de tráfego transportada pelos *lightpaths* que é roteada eletronicamente. Consequentemente, ele representa uma métrica mais acurada para o volume total de comutação/roteamento eletrônico realizado pela rede. Contudo, esse modelo só consegue capturar indiretamente os custos associados aos LTEs.

3. Número máximo de *lightpaths* incidentes a um elemento de rede:

Este modelo de custo é apropriado para se minimizar os custos associados aos LTEs e ao roteamento eletrônico do elemento de rede em que tais valores são máximos e não para a rede como um todo. Dessa forma, essa função objetivo (do tipo *min-max*) auxilia no estabelecimento de limites sobre os equipamentos (e, consequentemente, sobre o custo) de cada elemento da rede individualmente, o que pode vir a ser de grande interesse.

Inicialmente, neste trabalho, o **TGP** e, mais especificamente, o **TGP-MH** é explorado para uma rede ótica PPWDM (*Point-to-Point WDM* [55]) com topologia arbitrária em que, semelhante ao trabalho em [169], cada elemento da rede é um GOXC (isto é, possui capacidade de realizar *grooming*); contudo, nenhum limite sobre tal capacidade será inicialmente considerado, como em [149]. Mais adiante, são discutidos os ajustes necessários à formulação originalmente proposta para se capturar as situações em que a capacidade de realização de *grooming* esteja restrita a apenas alguns dos elementos da rede ou em que tal capacidade esteja limitada pelo número de *transceivers* (ou ADMs) existentes em cada elemento da rede.

Quanto à função objetivo, neste trabalho utiliza-se a primeira das funções discutidas acima (aquela em que se minimiza o número total de *lightpaths*), tendo em vista sua ampla adoção na maioria dos trabalhos da literatura, além de sua maior “qualidade” (se comparada a segunda opção) e simplicidade (em relação a terceira função). Mais especificamente, uma variação de tal função é apresentada em detalhes mais adiante (no final da subseção 3.1.2) e utilizada nos experimentos computacionais.

Mais ainda, este trabalho se concentra no estudo do **TGP** para redes óticas WDM com topologias arbitrárias em que não há *explicitamente* nenhum conversor de comprimento de onda nos elementos da rede, tendo em vista o elevado custo de tais conversores (fazendo com que na realidade raramente sejam adotados). Isso, porém, não impede que um tráfego (ou conjunto de tráfegos) seja retirado de um comprimento de onda por um elemento da rede e inserido novamente na rede, via roteamento eletrônico, utilizando outro comprimento de onda, desde que o “benefício” de tal operação seja compensado pelo seu alto custo.

3.1 Traffic Grooming Problem

Para se apresentar a formulação matemática do **TGP**, deve-se discutir primeiramente a adoção de uma representação estendida para a topologia da rede (subseção 3.1.1). Em seguida, a subseção 3.1.2 introduz a notação utilizada e a formulação do modelo de programação matemática propriamente dita. Essa formulação pode ser simplificada pela adoção de uma representação (grafo) em camadas, discutida na subseção 3.1.3, resultando em uma formulação condensada do problema (apresentada na subseção 3.1.4) que é utilizada no restante deste trabalho.

3.1.1 Representação Estendida

As informações básicas sobre a topologia de uma rede ótica (isto é, a disposição dos componentes óticos e de suas interconexões) podem ser representadas através de um grafo não direcionado, $H = (V, E)$, $v = |V|$, $e = |E|$. O tráfego t_p entre um par de nós do grafo H pode ser representado pela quádrupla (s_p, d_p, f_p, m_p) , em que s_p e d_p são, respectivamente, o nó de origem e o nó de destino do tráfego t_p , enquanto que f_p representa o volume de tráfego (em número de canais) que deve ser roteado entre s_p e d_p . O último elemento da quádrupla, m_p , representa, por sua vez, o número máximo de canais que podem ser transportados através de um comprimento de onda (ou *container virtual*¹). Pode-se, então, definir P como o conjunto de todos os tráfegos (doravante denominados produtos) existentes entre os nós de H , isto é, $P = \{t_p = (s_p, d_p, f_p, m_p) \mid s_p \in V, d_p \in V, f_p \in \mathbb{Z}_+, m_p \in \mathbb{Z}_+, 1 \leq p \leq n_p\}$, em que n_p é o número total de produtos. Vale dizer que P é, na verdade, um multiconjunto, isto é, P pode conter elementos iguais. Dessa forma, podem existir dois produtos distintos que compartilham a mesma origem e o mesmo destino, além de apresentarem volume e número máximo de canais iguais.

Apesar do grafo não direcionado H ser suficiente para se descrever as informações básicas sobre a topologia da rede ótica, ele não permite a representação dos componentes internos (multiplexadores e comutadores) dos elementos da rede. Tal representação se faz necessária para se modelar o **TGP**. Para tanto, cada nó da rede original será subdividido em vários outros, sendo representado na realidade por

¹Nas redes SDH/SONET, um *container virtual* representa um elemento básico de transporte e é constituído por um grupo de bytes alocados para transportar sinais de velocidade mais baixa e possui capacidade pré-estabelecida e fixa. Contudo, neste trabalho, apesar de se tornar “empresado” essa designação, sua capacidade é determinada em função da quantidade máxima de cada produto que um comprimento de onda é capaz de transportar (ver subseção seguinte).

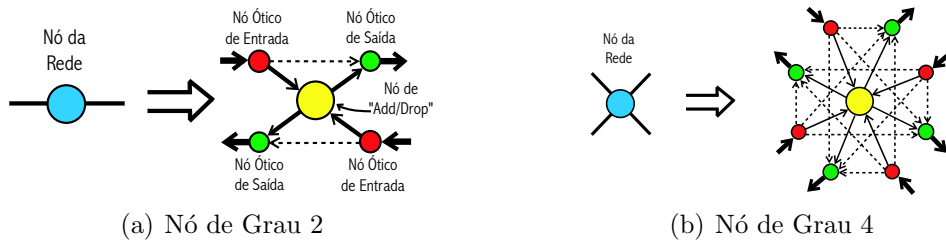


Figura 3.1: Subdivisão de um Nó

um conjunto de nós e arcos (ver Fig. 3.1(a)). Para cada nó da rede original haverá sempre um nó (denominado nó de *add/drop*) para representar o “mecanismo” de *add/drop* de cada elemento da rede, por exemplo, a parte elétrica/eletrônica do mesmo. Além disso, para cada porta de E/S de um nó da rede original, um par de nós será adicionado ao grafo (um representando a porta de entrada e o outro, a porta de saída). Estes nós são denominados nós óticos, sendo que todos os nós óticos que pertencem a um mesmo nó da rede original serão conectados ao nó de *add/drop* desse elemento da rede. Estas conexões são usadas para representar os multiplexadores de *add/drop* existentes no elemento em questão e estão representadas na Figura 3.1(a) através de linhas cheias saindo ou entrando no nó de *add-drop*.

Finalmente todos os nós óticos de entrada que pertencem ao mesmo elemento da rede também são conectados a todos os nós óticos de saída (com exceção daquele utilizado na representação da mesma porta de E/S) para se representar a capacidade de se rotear um comprimento de onda sem a necessidade de conversão O–E–O (ótico-eletró-ótica). Na Figura 3.1(a) estas conexões estão representadas através de linhas pontilhadas entre os nós óticos de entrada e saída.

Cumprе ressaltar que tal representação se torna mais complexa na medida em que o grau do nó da rede original aumenta. Isso pode ser facilmente observado comparando-se a Figura 3.1(b) que exibe a representação estendida de um nó de grau 4 com a Figura 3.1(a), em que o grau do nó original é apenas 2.

Seja g_i o grau do i -ésimo nó da rede original. Então o número de nós óticos, n_i^o , acrescidos à representação estendida do i -ésimo nó da rede original é $n_i^o = 2g_i$, enquanto que o número de arcos, a_i^o , adicionados para representar as conexões entre seus nós óticos e de seus nós óticos com o nó de *add/drop* é $a_i^o = g_i^2 + g_i$. Como $g_i = O(v)$ então $n_i^o = O(v)$ e $a_i^o = O(v^2)$, fazendo com que o número total de nós, n , da representação estendida seja igual a $n = v + \sum_{i \in V} n_i^o = O(v^2)$. Já o número total de arcos, a , da representação estendida é dado por $a = 2e + \sum_{i \in V} a_i^o = O(e + v^3)$, ou ainda, $a = O(v^3)$, uma vez que o número máximo de arcos de um grafo não direcionado é dado por $v(v-1)/2$, isto é, $e = O(v^2)$. Tais estimativas representam um aumento de uma ordem de grandeza tanto no número total de nós da representação estendida quanto em seu número total de arcos se comparada à rede original. Contudo tais estimativas representam o pior caso, ao passo que nas implementações reais de redes óticas é comum o emprego de topologias físicas em anel ou conjuntos (cadeias, árvores, etc...) de anéis, de modo a suportar eventuais falhas, fazendo com que o grau da maioria dos nós seja igual a 2, ou pouco maior que isto. Para o grafo H da Figura 3.2(a), obtém-se como resultado a representação estendida G , exibida na Figura 3.2(b) em que $n = |N| = 25$ e $a = |A| = 40$.

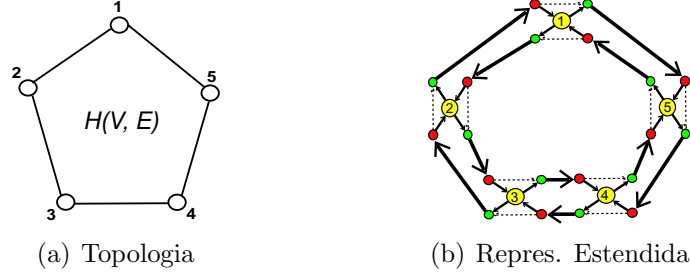


Figura 3.2: Exemplo da Representação Estendida

3.1.2 Formulação Matemática do TGP

Deve-se considerar a seguinte notação a ser utilizada na formulação de programação matemática associada ao **TGP** cuja representação estendida seja $G = (N, A)$:

- N representa o conjunto de nós, que pode ser particionado em dois subconjuntos disjuntos, $N = N^e \cup N^o$, em que N^e representa o conjunto de nós de *add-drop* e N^o representa o conjunto de nós óticos;
- A representa o conjunto de arcos, que pode ser particionado em dois subconjuntos disjuntos, $A = A^e \cup A^o$, em que A^e representa o conjunto de arcos entre nós de *add-drop* e os nós óticos, enquanto que A^o representa o conjunto de arcos entre nós óticos;
- C representa o conjunto de *containers virtuais* (ou comprimentos de onda), tal que $C = \{1, 2, \dots, W\}$, em que W representa o número total de *containers virtuais* (ou comprimentos de onda) disponíveis;
- P representa o conjunto de produtos a serem transportados, tal que $P = \{t_p = (s_p, d_p, f_p, m_p) \mid s_p \in N^e, d_p \in N^e, f_p \in \mathbb{Z}_+, m_p \in \mathbb{Z}_+, 1 \leq p \leq n_p\}$, em que n_p é o número total de produtos;
- s_p representa o nó de origem do produto $t_p \in P$;
- d_p representa o nó de destino do produto $t_p \in P$;
- f_p representa a quantidade (em número de canais) do produto $t_p \in P$ que deve ser transportada de s_p a d_p ;
- m_p representa a quantidade máxima (em número de canais) do produto $t_p \in P$ que pode ser transportada em um *container virtual* (ou comprimento de onda);
- c_{ijk} representa o custo de utilização do *container virtual* (ou comprimento de onda) $k \in C$ no arco $(i, j) \in A$.

Além disso, as seguintes variáveis são utilizadas na formulação do **TGP**:

- f_{ijk}^p representa a quantidade (em canais) do produto $t_p \in P$ que trafega através do arco $(i, j) \in A$ utilizando o *container virtual* $k \in C$;
- w_{ijk} indica o uso do *container virtual* $k \in C$ no arco $(i, j) \in A$ no transporte de algum produto.

Por fim, a capacidade de um *container virtual* (ou comprimento de onda) é dada por:

$$\Delta = \text{mmc}^2 \{ m_p \mid \forall t_p = (s_p, d_p, f_p, m_p) \in P \}. \quad (3.1)$$

²Mínimo múltiplo comum.

Ao passo que a contribuição (ou melhor, a ocupação) correspondente a uma unidade de fluxo de um produto transportado em um *container virtual* é representada por:

$$\delta_p = \Delta/m_p, \quad \forall t_p \in P. \quad (3.2)$$

Sendo assim, a formulação de programação matemática, *TGP*, associada ao problema é dada por:

$$(TGP) \quad \min \sum_{k \in C} \sum_{(i,j) \in A} c_{ijk} w_{ijk} \quad (3.3a)$$

sujeito a:

$$\sum_{k \in C} \sum_{(i,j) \in A^e} f_{ijk}^p - \sum_{k \in C} \sum_{(j,i) \in A^e} f_{jik}^p = \begin{cases} f_p, i = s_p \\ -f_p, i = d_p \\ 0, i \neq s_p \neq d_p \end{cases}, \forall t_p \in P, \forall i \in N^e \quad (3.3b)$$

$$\sum_{(i,j) \in A} f_{ijk}^p - \sum_{(j,i) \in A} f_{jik}^p = 0, \quad \forall t_p \in P, \forall i \in N^o, \forall k \in C \quad (3.3c)$$

$$\sum_{t_p \in P} \delta_p f_{ijk}^p \leq \Delta w_{ijk}, \quad \forall (i,j) \in A, \forall k \in C \quad (3.3d)$$

$$\sum_{(i,j) \in A} w_{ijk} - \sum_{(j,i) \in A} w_{jik} = 0, \quad \forall i \in N^o, \forall k \in C \quad (3.3e)$$

$$f_{ijk}^p \geq 0, \quad \forall t_p \in P, \forall (i,j) \in A, \forall k \in C \quad (3.3f)$$

$$w_{ijk} \in \{0, 1\}, \quad \forall (i,j) \in A, \forall k \in C \quad (3.3g)$$

$$f_{ijk}^p \text{ inteiro}, \quad \forall t_p \in P, \forall (i,j) \in A, \forall k \in C \quad (3.3h)$$

A função objetivo dada por (3.3a) procura minimizar o custo total de utilização/alocação dos “containers virtuais”. As restrições (3.3b) e (3.3c) garantem a conservação de fluxo dos produtos para os nós de *add-drop* e para os nós óticos. Vale ressaltar que nas restrições (3.3b) todo o fluxo que entra ou sai de um nó de *add-drop* (independentemente do comprimento de onda utilizado para transportá-lo) é considerado nos somatórios, permitindo assim que o fluxo de um produto, que entre em um nó de *add-drop* utilizando um dado comprimento de onda, venha a deixar o mesmo nó através de um outro comprimento de onda. Já o mesmo não se aplica aos nós óticos, uma vez que as restrições (3.3c) são escritas separadamente para cada comprimento de onda $k \in C$.

As restrições (3.3d) impõem um limite sobre o volume total de produtos transportados através de um *container virtual* no arco $(i,j) \in A$, isto é, um limite sobre a capacidade dos *containers virtuais*, além de estabelecer uma ligação entre as variáveis de fluxo (f_{ijk}^p) e de decisão (w_{ijk}). Já as restrições (3.3e) garantem a continuidade de comprimento de onda dos *lightpaths*, além de serem responsáveis, juntamente com a estrutura do grafo, pela não bifurcação dos *lightpaths*. Finalmente, as restrições

(3.3f), (3.3g) e (3.3h) definem as variáveis de fluxo (f_{ijk}^p) como variáveis inteiras e não negativas, enquanto que as variáveis de decisão (w_{ijk}) são definidas como binárias.

Deve-se mencionar que essa formulação (*TGP*) possui $aW(n_p + 1)$ variáveis relacionadas através de $n_p n_e + aW + n_o W(n_p + 1)$ restrições, em que $n_e = |N^e|$ e $n_o = |N^o|$ e não computando as restrições (3.3f)–(3.3h) que definem as variáveis.

Detalhamento da Função Objetivo

A função objetivo apresentada anteriormente (3.3a) permite associar um custo diferenciado c_{ijk} para utilização (alocação) de cada *container virtual* (ou comprimento de onda) $k \in C$ em cada um dos arcos $(i, j) \in A$.

Como descrito anteriormente, neste trabalho utiliza-se a função que minimiza o número total de *lightpaths*, representado por N_L . Sendo assim, basta associar aos arcos incidentes aos nós de *add-drop* um custo igual a $1/2$; enquanto que os demais custos serão iguais a zero, isto é:

$$c_{ijk} = \begin{cases} \frac{1}{2}, & \text{se } i \in N^e \text{ ou } j \in N^e; \\ 0, & \text{caso contrário;} \end{cases} \quad \forall k \in C, \forall (i, j) \in A. \quad (3.4)$$

Como cada *lightpath* contém apenas dois arcos incidentes a nós de *add-drop* (um na sua origem e outro em seu destino), pode-se concluir, para uma dada solução (\bar{w}, \bar{f}) de *TGP*, que:

$$N_L = \sum_{k \in C} \sum_{(i, j) \in A} c_{ijk} \bar{w}_{ijk}. \quad (3.5)$$

Além disso, testes preliminares indicaram que se podia adotar um outro padrão para os custos c_{ijk} , em que os custos de arcos incidentes a nós de *add-drop* sejam cerca de 2 ou 3 ordens de grandeza maiores que aqueles associados aos demais arcos. Dessa maneira, pode-se direcionar o processo de busca de soluções para se obter uma solução não apenas com o número mínimo de *lightpaths* mas também cujos *lightpaths* possuam o menor comprimento possível (em relação aos custos adotados). Poder-se-ia, por exemplo, assumir que:

$$c_{ijk} = \begin{cases} \tilde{c}, & \text{se } (i, j) \in A^e; \\ c_1 l_{ij} + c_2, & \text{se } (i, j) \in \bar{A}^o; \\ 0, & \text{caso contrário;} \end{cases} \quad \forall k \in C, \forall (i, j) \in A, \quad (3.6)$$

em que \bar{A}^o representa o conjunto de arcos entre nós óticos de diferentes elementos da rede, l_{ij} representa o comprimento “real” do enlace físico utilizado na ligação entre os elementos da rede, c_1 e c_2 representam as parcelas variável e fixa, respectivamente, associadas ao estabelecimento de enlace físico (por exemplo, custo por quilômetro da fibra e dos equipamentos/placas necessários em suas extremidades) e $\tilde{c} = 100 \times \max \{ c_{LTE}, \max_{(i, j) \in \bar{A}^o} \{ c_1 l_{ij} + c_2 \} \}$, em que c_{LTE} representa o custo associado a instalação de um LTE.

Por fim, pode-se, ainda, adotar uma versão mais simples em que:

$$c_{ijk} = \begin{cases} 1000 + k, & \text{se } (i, j) \in A^e; \\ 10 + k, & \text{se } (i, j) \in A^o; \end{cases} \quad \forall k \in C, \forall (i, j) \in A. \quad (3.7)$$

Nesta última versão, além de se direcionar (de forma simplificada) o processo de resolução para obtenção de soluções cujos *lightpaths* possuam o menor comprimento possível, a inclusão de k faz com que os *containers virtuais* (ou comprimentos de onda) de menor número sejam utilizados preferencialmente. Seu efeito final é reduzir o número de soluções alternativas, o que eventualmente pode auxiliar o algoritmo de resolução.

3.1.3 Representação em Camadas

Nesta subseção apresenta-se um grafo em camadas obtido a partir da representação estendida discutida anteriormente (ver subseção 3.1.1). Esse grafo em camadas é utilizado para se reformular o problema obtendo um modelo mais simples e compacto para o **TGP**. Tal modelo simplificado pode ser visto como uma extensão dos modelos apresentados na literatura para o problema de *network loading* [11, 102] e, mais especificamente, para o problema de *minimum cost capacity installation* [20].

As principais razões que motivaram a adoção dessa representação em camadas são duas. Em primeiro lugar, diferentemente das referências mencionadas no parágrafo anterior, considerou-se neste trabalho que o número de facilidades disponíveis (isto é, *containers virtuais* ou comprimentos de onda em cada arco) já foi previamente estabelecido e não está sujeito a otimização pelo modelo proposto, semelhante ao adotado em [45]. Tal suposição não é absurda uma vez que em muitas situações reais de planejamento é comum se deparar com tal tipo de restrição. Se esse não for o caso, deve-se ter em mente que há uma profusão de referências na literatura, a exemplo dos trabalhos em [142, 154, 164], que pode ser utilizada para se limitar inferior e superiormente o número de facilidades necessárias a uma instância específica do **TGP**. Esses limites poderiam então servir de base para o desenvolvimento de um procedimento de busca que determinasse o número ótimo de facilidades necessárias, utilizando-se para tanto de um algoritmo que solucione o problema para um dado número fixo de facilidades.

Outra razão para a adoção de uma representação em camadas está ligada à simplificação de diversos aspectos tanto da formulação quanto dos procedimentos de resolução propostos mais adiante como, por exemplo, a utilização de implementações mais simples e eficientes para o cálculo de caminhos mais curtos.

Neste grafo em camadas cada *container virtual* (ou comprimento de onda) é representado por uma camada distinta. Os nós óticos da representação estendida são duplicados em cada camada, bem como as conexões (arcos) entre eles. Já cada um dos nós de *add-drop*, por sua vez, é mantido inalterado sem, contudo, estar associado a nenhuma camada em particular. As conexões entre estes e as duplicatas dos nós óticos são estabelecidas da seguinte forma: para cada arco saindo do nó de *add-drop* i e entrando no nó ótico j da representação estendida é adicionado um arco distinto entre o correspondente nó de *add-drop* da representação em camadas e cada uma das duplicatas do nó ótico j . De forma análoga, para cada arco saindo do nó ótico j e entrando no nó de *add-drop* i da representação estendida é adicionado um arco distinto entre cada uma das duplicatas do nó ótico j e o correspondente nó de *add-drop* da representação em camadas.

Para a representação inicial da topologia — grafo não direcionado — de uma

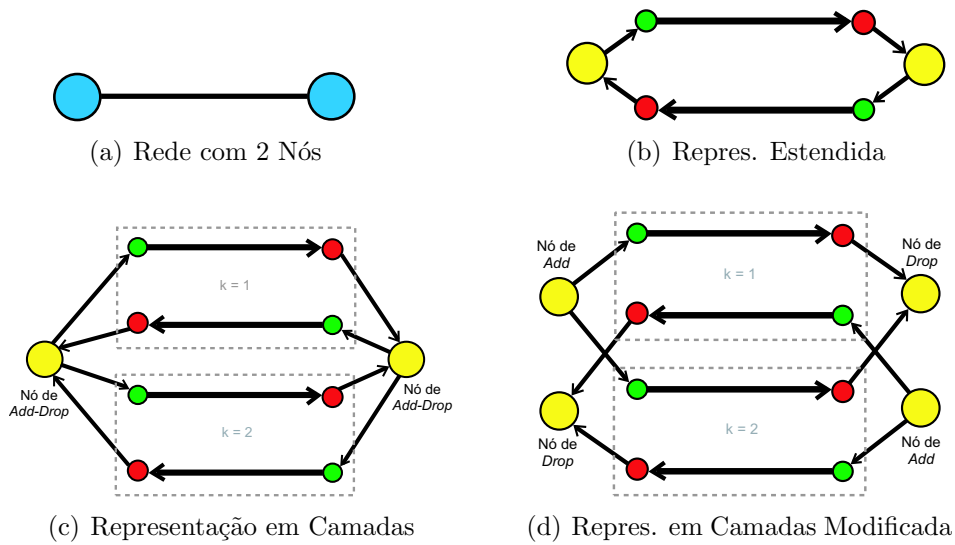


Figura 3.3: Representações de uma Rede com 2 Nós

rede com apenas dois nós exibida na Figura 3.3(a), obtém-se, após a adição dos nós óticos e de suas conexões, conforme descrito na subseção 3.1.1, a representação estendida mostrada na Figura 3.3(b). Caso o número total de *containers virtuais* (ou comprimentos de onda) disponíveis seja igual a 2, deve-se duplicar o conjunto de nós óticos, bem como suas interconexões, de modo a se obter uma representação em camadas a partir da representação estendida. Além disso, deve-se realizar a conexão de cada nó de *add-drop* com os respectivos nós óticos em cada camada, como descrito anteriormente. A Figura 3.3(c) exibe a representação em camadas correspondente a representação estendida da Figura 3.3(b) considerando um disponibilidade de 2 *containers virtuais*. Em tal figura, os nós óticos pertencentes a mesma camada, isto é, que se utilizam do mesmo *container virtual*, encontram-se agrupados através de um retângulo pontilhado.

Durante a descrição dos procedimentos de resolução uma versão modificada desta representação em camadas também é utilizada. Nela, (ver Figura 3.3(d)) cada um dos nós de *add-drop* é substituído por um par de nós representando, assim, separadamente, a fonte (responsável pelo *add*) e o sumidouro (responsável pelo *drop*) associada ao nó de *add-drop* original. Esses novos nós continuam não sendo associados a nenhuma camada em particular, ao passo que as conexões entre eles e os nós óticos serão estabelecidas da seguinte maneira: para cada arco saindo do nó de *add-drop* i e entrando no nó ótico j da representação estendida é adicionado um arco distinto entre o recém criado nó de *add* e cada uma das duplicatas do nó ótico j . De forma análoga, para cada arco saindo do nó ótico j e entrando no nó de *add-drop* i da representação estendida é adicionado um arco distinto entre cada uma das duplicatas do nó ótico j e o recém criado nó de *drop*.

A real necessidade dessa representação em camadas modificada se fará notar na medida que, durante os procedimentos de resolução, se fizer necessário representar tanto a quantidade de *containers virtuais* que devem sair de um nó de *add-drop* quanto a quantidade de *containers virtuais* que devem chegar no mesmo nó. A principal vantagem da representação em separado das partes responsáveis pela funções

de *add* e de *drop* de produtos está em se poder associar a oferta de *containers virtuais* ao nó de *add*, enquanto que a demanda de *containers virtuais* é associada ao nó de *drop*. Esse procedimento de representação simultânea da oferta e da demanda de *containers virtuais* não seria possível se apenas um nó fosse utilizado. Maiores detalhes sobre a utilização da representação em camadas modificada são descritos juntamente com os procedimentos de resolução.

3.1.4 Formulação Matemática Condensada do TGP

A adoção da representação em camadas permite reformular o problema de forma mais simples e compacta. Isso se deve principalmente à “eliminação”, melhor seria, substituição de um dos parâmetros utilizados na definição original do problema.

Em verdade, os modelos apresentados na subseção 3.1.2, juntamente com a avaliação de suas dimensões feita ao final da mesma subseção, deixam claro a relação existente entre a complexidade da formulação do problema e a quantidade de nós, arcos, produtos e *containers virtuais* (ou comprimentos de onda) que são utilizados na representação do **TGP**. O uso da representação em camadas elimina da formulação do problema toda e qualquer referência explícita ao conjunto de *containers virtuais*. Para tanto, cumpre ressaltar a necessidade de um aumento das quantidades de nós e arcos que passam a ser utilizados na formulação em substituição aos *containers virtuais*.

Essa representação implícita dos *containers virtuais* não conduz diretamente — pelo menos em um primeiro momento — a uma redução da complexidade do problema nem de suas dimensões (não se tratando, por exemplo, de uma operação de projeção de um conjunto de variáveis em outro). Contudo, ela propicia uma simplificação real da notação e, conseqüentemente, um melhor entendimento da formulação, bem como do próprio problema.

Deve-se considerar a seguinte notação a ser utilizada na formulação matemática condensada para o **TGP** cuja representação em camadas, conforme descrito na subseção 3.1.3, seja dada pelo grafo $G_\ell = (N_\ell, A_\ell)$:

- N_ℓ representa o conjunto de nós da representação em camadas, que pode ser particionado em dois subconjuntos disjuntos, $N = N_\ell^e \cup N_\ell^o$, em que N_ℓ^e representa o conjunto de nós de *add-drop* e N_ℓ^o representa o conjunto de nós óticos independentemente das camadas;
- A_ℓ representa o conjunto de arcos da representação em camadas;
- c_ℓ representa o custo de utilização do arco $\ell \in A_\ell$.

As tuplas do conjunto de produtos devem ser ajustadas de modo que os nós de origem e destino de produtos correspondam aos nós pertencentes a N_ℓ^e .

Além disso, as seguintes variáveis são utilizadas na formulação condensada do **TGP**:

- f_ℓ^p representa a quantidade (em canais) do produto $t_p \in P$ que trafega através do arco $\ell \in A_\ell$;
- w_ℓ indica o uso do arco $\ell \in A_\ell$ no transporte de algum produto.

Mais ainda, a capacidade máxima de qualquer um dos arcos é dada por Δ (a mesma capacidade máxima atribuída antes a um *container virtual*) e cada unidade de fluxo de um produto continua a consumir δ_p unidades de capacidade de transporte.

Por fim, $A_\ell^+(i)$ é utilizado para representar o conjunto de todos arcos que saem de um nó $i \in N_\ell$, enquanto que $A_\ell^-(i)$ representa o conjunto de todos os arcos que chegam em um nó $i \in N_\ell$, isto é:

$$A_\ell^+(i) = \{\ell \in A_\ell \mid \ell = (i, j), j \in N_\ell\}, \forall i \in N_\ell; \quad (3.8)$$

$$A_\ell^-(i) = \{\ell \in A_\ell \mid \ell = (j, i), j \in N_\ell\}, \forall i \in N_\ell. \quad (3.9)$$

Dessa forma, uma formulação condensada, $TGPC$, associada ao problema é dada por:

$$(TGPC) \quad \min \sum_{\ell \in A_\ell} c_\ell w_\ell \quad (3.10a)$$

sujeito a:

$$\sum_{\ell \in A_\ell^+(i)} f_\ell^p - \sum_{\ell \in A_\ell^-(i)} f_\ell^p = b_i^p, \quad \forall t_p \in P, \forall i \in N_\ell \quad (3.10b)$$

$$\sum_{\ell \in A_\ell^+(i)} w_\ell - \sum_{\ell \in A_\ell^-(i)} w_\ell = 0, \quad \forall i \in N_\ell^o \quad (3.10c)$$

$$\sum_{t_p \in P} \delta_p f_\ell^p \leq \Delta w_\ell, \quad \forall \ell \in A_\ell \quad (3.10d)$$

$$f_\ell^p \geq 0, \quad \forall t_p \in P, \forall \ell \in A_\ell \quad (3.10e)$$

$$w_\ell \in \{0, 1\}, \quad \forall \ell \in A_\ell \quad (3.10f)$$

$$f_\ell^p \text{ inteiro}, \quad \forall t_p \in P, \forall \ell \in A_\ell \quad (3.10g)$$

em que b_i^p é dado por:

$$b_i^p = \begin{cases} f_p & , \text{ se } i = s_p \\ -f_p & , \text{ se } i = d_p \\ 0 & , \text{ se } i \neq s_p \neq d_p \end{cases} \quad \forall t_p \in P, \forall i \in N_\ell. \quad (3.11)$$

Cumpre reafirmar que a formulação $TGPC$ dada por (3.10a)–(3.10g) é totalmente equivalente àquela apresentada anteriormente, TGP , através das equações (3.3a)–(3.3h). A função objetivo dada por (3.10a) procura minimizar o custo total de utilização/alocação dos arcos da representação em camadas. As restrições (3.10b) garantem a conservação de fluxo dos produtos para todos os nós da representação em camadas. Já as restrições (3.10c) garantem que a quantidade de arcos alocados que saem de um nó ótico é igual a quantidade de arcos alocados que chegam no mesmo nó (sendo, assim, semelhantes às restrições (3.3e), responsáveis pela continuidade dos *lightpaths*). As restrições (3.10d), por sua vez, impõem um limite sobre o volume total de produtos transportados através de um arco $\ell \in A_\ell$, além de estabelecer uma ligação entre as variáveis de fluxo (f_ℓ^p) e de decisão (w_ℓ). Finalmente, as restrições (3.10e), (3.10f) e (3.10g) definem as variáveis de fluxo (f_ℓ^p) como variáveis inteiras e não negativas, enquanto que as variáveis de decisão (w_ℓ) são definidas como binárias.

Quanto ao número de variáveis, é fácil constatar que a formulação condensada (TGP_C) possui $a_\ell(n_p + 1)$ variáveis, em que $a_\ell = |A_\ell|$. Contudo, de acordo com o procedimento descrito para a construção da representação em camadas, sabe-se que $a_\ell = aW$ e, conseqüentemente, TGP_C possui $aW(n_p + 1)$ variáveis. Já em relação às restrições, essa formulação possui $n_p n_\ell + n_\ell^o + a_\ell$ restrições, em que $n_\ell = |N_\ell|$ e $n_\ell^o = |N_\ell^o|$. Pela descrição do processo de construção da representação em camadas é fácil notar que $n_l = n_e + n_\ell^o$ e $n_\ell^o = n_o W$. Logo o número de restrições de TGP_C pode ser reescrito como $n_p(n_e + n_o W) + n_o W + aW$, ou ainda, $n_p n_e + aW + n_o W(n_p + 1)$. Ao se comparar tais valores com aqueles apresentados anteriormente para a formulação original (no final da subseção 3.1.2) comprova-se que, na realidade, ambas as formulações possuem o mesmo número de variáveis e restrições.

Além dessa nova formulação, TGP_C , implicar em uma óbvia simplificação do conjunto de restrições, ela possibilita identificar mais facilmente a relação existente entre o **TGP** e outros problemas de projeto/planejamento de redes capacitadas encontrados na literatura [11, 20, 102]. Tais problemas têm recebido uma considerável atenção nesta última década em virtude, principalmente, da ampliação do uso de redes óticas como tecnologia de transporte de alta capacidade.

Sob esse enfoque o **TGP** poderia ser descrito como o problema de se instalar facilidades de transporte (capacidade) nos arcos de uma rede, de modo a possibilitar o roteamento de todo o tráfego de produtos existentes a um custo total (de instalação) mínimo. A essa descrição deve-se acrescentar dois fatos que tornam o **TGP** distinto dos demais problemas de projeto/planejamento de redes capacitadas encontrados até este momento na literatura.

Em primeiro lugar, a formulação do **TGP** deve incorporar as limitações funcionais da tecnologia existente. Isso corresponde, na realidade, à inclusão de restrições adicionais, quer seja sobre o roteamento, quer seja sobre a instalação de facilidades. No caso específico do **TGP**, tais restrições são de caráter estrutural na medida em que devem limitar a estrutura (topologia) criada pela instalação de facilidades. Na formulação condensada (TGP_C) as restrições (3.10c) cumprem esse papel ao garantir que, para cada facilidade instalada em um arco que chega em um nó ótico, deva existir uma facilidade correspondente instalada em um arco que sai do mesmo nó. Tais restrições, juntamente com o fato de, por construção, um nó ótico possuir grau de entrada (ou de saída) sempre igual a 1, são responsáveis pela *continuidade* de uso da mesma facilidade em tal nó da rede. Em outras palavras, um conjunto de produtos transportado até um nó ótico utilizando determinada facilidade deve continuar a ser transportado através do uso da mesma facilidade só sendo permitida sua substituição por outra quando esse conjunto de produtos alcançar um nó de *add-drop* com capacidade para a realização de *grooming*. Mais ainda, a facilidade utilizada no transporte de produtos não pode ser “subdividida” em um nó ótico, evitando-se assim a “bifurcação” do *lightpath*. Em redes óticas WDM esse tipo de limitação é muito comum, sendo conhecida como restrição de continuidade de comprimento de onda (*wavelength continuity constraint*) [43].

Além da inclusão de tais restrições estruturais, a formulação do **TGP** também se diferencia das demais pela forma de se contabilizar a ocupação da capacidade instalada pelos produtos. Tradicionalmente, uma unidade de fluxo de qualquer produto

é considerada equivalente a uma unidade de capacidade instalada. Apesar de ser comum a representação de diferentes tipos de facilidades (com diferentes capacidades), quase que a totalidade dos trabalhos encontrados até então na literatura tem assumido essa hipótese de equivalência. No caso da formulação aqui apresentada para o **TGP**, as restrições (3.10d) permitem que os produtos contribuam de forma diferenciada para a ocupação da capacidade instalada.

Isto poderia ser contornado através da substituição das variáveis de fluxo f_ℓ^p pelas variáveis $x_\ell^p \in \mathbb{Z}, \forall t_p \in P, \forall \ell \in A_\ell$, tais que $x_\ell^p = \delta_p f_\ell^p, \forall t_p \in P, \forall \ell \in A_\ell$. Essa transformação pode ser entendida como a adoção do conceito de um “produto equivalente”. Sendo assim, o problema original poderia ser resolvido a partir da transformação das ofertas e demandas de seus produtos (f_p) para ofertas e demandas dos “produtos equivalentes” ($\delta_p f_p$). Uma vez resolvido o problema para os “produtos equivalentes”, bastaria realizar a conversão da solução encontrada (em particular, do roteamento de “produtos equivalentes” para o roteamento dos produtos do problema original).

Infelizmente, mesmo que se obtenha uma solução inteira para o problema envolvendo os “produtos equivalentes” (o que já não é simples), tal solução não corresponderia necessariamente a uma solução do problema original, tendo em vista que os valores de fluxo de qualquer “produto equivalente” a um produto original $t_p \in P$ não seriam necessariamente múltiplos inteiros de δ_p . Através da adição de novas restrições ao problema, poder-se-ia garantir as condições necessárias para que uma solução envolvendo “produtos equivalentes” corresponda a exatamente uma solução do problema original. Contudo, em vista do que foi exposto, a formulação original teria de ser alterada e, portanto, seria distinta das demais encontradas até este momento na literatura.

Apesar da substituição de variáveis propostas não auxiliar na simplificação (e resolução) do problema, um artifício semelhante (só que com variáveis reais) será utilizado no estudo da estrutura do poliedro associado ao conjunto de soluções do **TGP** e, como será visto mais adiante, mostrar-se-á de grande ajuda.

3.2 Traffic Grooming Esparso

Como destacado nos objetivos deste trabalho (ver seção 1.2) e reafirmado na introdução deste capítulo, são discutidos os ajustes necessários à formulação originalmente proposta para se capturar as situações em que a capacidade de realização de *grooming* está limitada a apenas alguns dos elementos da rede. Esta seção, portanto, apresenta tais ajustes.

3.2.1 Limitando o Número de GOXCs

Para se limitar a capacidade de *grooming* a um subconjunto dos nós de *add-drop* (isto é, reduzir o número GOXCs na rede), deve-se particionar tal conjunto em dois subconjuntos, $N^e = N^g \cup N^{ad}$, em que N^g representa o conjunto de nós que realizam *grooming* além de realizarem *add-drop* e N^{ad} representa o conjunto de nós que realizam apenas *add-drop*.

A escolha de quais nós pertencem a N^g pode ser orientada por diversos parâmetros. Como exemplo, em uma rede formada por vários anéis interconectados, uma escolha sensata seria alocar tal capacidade de realização de *grooming* aos nós de interconexão entre os anéis, a semelhança do que, via de regra, é realizado pelas operadoras de rede em projetos reais.

Considere ainda o conjunto de nós que realizam apenas *add-drop* e que também são nós de oferta e/ou demanda de um produto $t_p \in P$; pode-se, então, definir $N_p^{ad} = \{i \in N^{ad} \mid i = s_p \text{ ou } i = d_p\}$.

Além dessa subdivisão do conjunto de nós de *add-drop*, as seguintes variáveis são necessárias à formulação desta versão limitada do **TGP** – *Sparse Traffic Grooming Problem* ou, simplesmente, **STGP** – e devem ser acrescentadas àquelas definidas anteriormente, na subseção 3.1.2:

f_{ik}^p representa a quantidade do produto $t_p \in P$ inserida/retirada no nó $i \in N^{ad}$ utilizando o *container virtual* $k \in C$.

Sendo assim, a formulação de programação matemática, *STGP*, associada a essa versão do problema é dada por:

$$(STGP) \quad \min \sum_{k \in C} \sum_{(i,j) \in A} c_{ijk} w_{ijk} \quad (3.12a)$$

sujeito a:

$$\sum_{k \in C} \sum_{(i,j) \in A^e} f_{ijk}^p - \sum_{k \in C} \sum_{(j,i) \in A^e} f_{jik}^p = \begin{cases} f_p, i = s_p \\ -f_p, i = d_p \\ 0, i \neq s_p \neq d_p \end{cases}, \forall t_p \in P, \forall i \in N^g \quad (3.12b)$$

$$\sum_{(i,j) \in A^e} f_{ijk}^p - \sum_{(j,i) \in A^e} f_{jik}^p = \begin{cases} f_{ik}^p, i = s_p \\ -f_{ik}^p, i = d_p \\ 0, i \neq s_p \neq d_p \end{cases}, \forall t_p \in P, \forall i \in N^{ad}, \forall k \in C \quad (3.12c)$$

$$\sum_{k \in C} f_{ik}^p = f_p, \forall t_p \in P, \forall i \in N_p^{ad} \quad (3.12d)$$

$$\sum_{(i,j) \in A} f_{ijk}^p - \sum_{(j,i) \in A} f_{jik}^p = 0, \forall t_p \in P, \forall i \in N^o, \forall k \in C \quad (3.12e)$$

$$\sum_{t_p \in P} \delta_p f_{ijk}^p \leq \Delta w_{ijk}, \forall (i,j) \in A, \forall k \in C \quad (3.12f)$$

$$\sum_{(i,j) \in A} w_{ijk} - \sum_{(j,i) \in A} w_{jik} = 0, \forall i \in N^o, \forall k \in C \quad (3.12g)$$

$$f_{ijk}^p \geq 0, \forall t_p \in P, \forall (i,j) \in A, \forall k \in C \quad (3.12h)$$

$$f_{ik}^p \geq 0, \forall t_p \in P, \forall i \in N_p^{ad}, \forall k \in C \quad (3.12i)$$

$$w_{ijk} \in \{0, 1\}, \forall (i,j) \in A, \forall k \in C \quad (3.12j)$$

$$f_{ijk}^p \text{ inteiro}, \forall t_p \in P, \forall (i,j) \in A, \forall k \in C \quad (3.12k)$$

$$f_{ik}^p \text{ inteiro}, \forall t_p \in P, \forall i \in N_p^{ad}, \forall k \in C \quad (3.12l)$$

Nesta formulação, tanto a função objetivo (3.12a) como as restrições (3.12b)–(3.12l) são muito similares àquelas apresentadas pela formulação anterior. Vale

destacar que as restrições (3.12c) garantem a conservação de fluxo nos nós que realizam apenas *add-drop*. Enquanto que as restrições (3.12d) asseguram que a quantidade de produto inserida/retirada da rede através dos nós que realizam apenas *add-drop* é equivalente a oferta/demanda de tal produto.

A formulação apresentada poderia ser simplificada tendo em vista que um nó de *add-drop* só é capaz de inserir (ou retirar) os produtos que se originem (ou destinem) a ele, sendo vetado a ele qualquer manipulação dos demais produtos. Assim, as restrições (3.12c) seriam reescritas da seguinte forma:

$$\sum_{(s_p, j) \in A^e} f_{s_p j k}^p = f_{s_p k}^p, \forall t_p \in P, \forall k \in C, \text{ se } s_p \in N^{ad}; \quad (3.13a)$$

$$\sum_{(j, d_p) \in A^e} f_{j d_p k}^p = f_{d_p k}^p, \forall t_p \in P, \forall k \in C \text{ se } d_p \in N^{ad}. \quad (3.13b)$$

Cabe notar que esta formulação é totalmente compatível com a anterior, podendo-se, inclusive, admitir a primeira formulação (*TGP*) como um caso especial da segunda (*STGP*) pois, caso se atribua a capacidade de realizar *grooming* a todos os nós de *add-drop*, isto é, $N^g = N^e$ e $N^{ad} = \emptyset$, a segunda formulação (*STGP*) se reduz àquela apresentada anteriormente (*TGP*), uma vez que as restrições (3.12c) e (3.12d) serão eliminadas.

Para essa formulação (*STGP*), a quantidade de variáveis passa a depender também do número total de vezes em que um nó que realiza apenas *add-drop* também representa a fonte (nó de oferta) ou o sumidouro (nó de demanda) de um produto, $n_p^{ad} = \sum_{t_p \in P} |N_p^{ad}|$, resultando em um acréscimo de $W n_p^{ad}$ variáveis, ou ainda, $aW(n_p + 1) + W n_p^{ad}$ variáveis no total. Da mesma forma, haverá um acréscimo de $n_p n_{ad}(W - 1) + n_p^{ad}$ restrições, em que $n_{ad} = |N^{ad}|$, totalizando $n_p n_e + aW + n_o W(n_p + 1) + n_p n_{ad}(W - 1) + n_p^{ad}$ restrições.

3.2.2 Alterando a Representação em Camadas

A princípio não há grandes diferenças entre a representação em camadas descrita anteriormente (ver subseção 3.1.3) e aquela a ser adotada na representação de uma rede em que a capacidade de realização de *grooming* é limitada a um subconjunto dos elementos da rede. Sendo assim, a Figura 3.3(c) pode ser associada ao cenário em que todos os elementos possuem capacidade de realizar *grooming*. De forma a se eliminar a capacidade de realizar *grooming* de um elemento específico da rede, deve-se substituir o nó de *add-drop* correspondente a ele por um par de nós, conforme procedimento descrito ao final da subseção 3.1.3. Portanto, a Figura 3.3(d) representa a situação em que nenhum dos elementos da rede é capaz de realizar *grooming*.

Uma representação em camadas associada a uma instância do **STGP** deve estar entre esses dois extremos, em que os nós pertencentes a N^g (isto é, com capacidade de realizar *grooming*) são representados como os nós da Figura 3.3(c), enquanto que os demais (sem capacidade de realizar *grooming*) são representados com os nós da Figura 3.3(d).

Ao invés de se utilizar duas representações distintas para os nós, parece ser mais razoável representar cada elemento da rede utilizando a mesma abordagem. Sendo

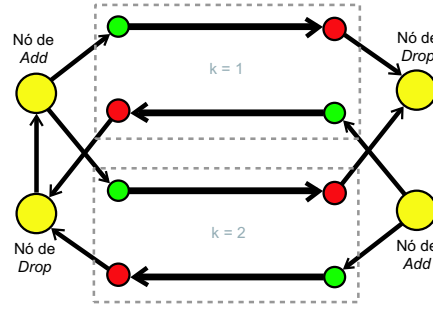


Figura 3.4: Representação em Camadas para **STGP**

assim, cada elemento da rede deve ser representado através de um par de nós como na Figura 3.3(d) e para cada um deles em que se possa realizar *grooming* irá existir um arco entre o seu nó de *drop* e o correspondente nó de *add*. A Figura 3.4 exhibe um exemplo de tal representação. O par de nós conectados (um de *add* e o outro de *drop*) à esquerda da figura representam um elemento com capacidade de realizar *grooming*, enquanto que o par não conectado à direita representa um elemento da rede sem tal capacidade.

Uma vez que se adote a representação descrita acima para as instâncias de **STGP**, a formulação condensada de programação matemática associada ao problema será praticamente a mesma dada anteriormente, isto é, TGP_C (ver subseção 3.1.4). Contudo, os arcos entre os nós de *drop* e de *add* devem ter custos nulos e não devem possuir nenhum limite sobre sua capacidade de fluxo, ou melhor, as restrições (3.10d) não devem ser aplicadas a eles.

Apesar da semelhança entre tais formulações, de forma a tornar evidente em que momento essa representação em camadas alternativa está ou não sendo utilizada, denominar-se-á de $STGP_C$ o uso da formulação condensada dada anteriormente associada a representação descrita acima para uma instância de **STGP**.

3.3 Limitando o Número de *Transceivers*

Na introdução deste capítulo, destacou-se também que seriam discutidos os ajustes necessários à formulação originalmente proposta para representar as situações em que a capacidade de realização de *grooming* é limitada pelo número de *transceivers* (ou ADMs) existentes em cada elemento da rede.

Desse modo, caso se deseje impor uma limitação sobre número de *transceivers* existentes em cada elemento da rede, deve-se acrescentar o seguinte conjunto de restrições à formulação:

$$\sum_{k \in C} \sum_{(i,j) \in A^e} w_{ijk} \leq N_i^T, \forall i \in N^e; \quad (3.14)$$

$$\sum_{k \in C} \sum_{(j,i) \in A^e} w_{jik} \leq N_i^R, \forall i \in N^e. \quad (3.15)$$

em que N_i^T e N_i^R representam o número de transmissores e receptores óticos, respectivamente, existentes no elemento da rede associado ao nó de *add-drop* $i \in N^e$.

Tais restrições podem ser utilizadas tanto em *TGP* como em *STGP*, porém, para utilizá-las em conjunto com as formulações condensadas, *TGP_C* e *STGP_C*, torna-se necessário reescrevê-las da seguinte forma:

$$\sum_{\ell \in A_{\ell}^{+}(i)} w_{\ell} \leq N_i^T, \forall i \in N_{\ell}^e; \quad (3.16)$$

$$\sum_{\ell \in A_{\ell}^{-}(i)} w_{\ell} \leq N_i^R, \forall i \in N_{\ell}^e. \quad (3.17)$$

3.4 *Traffic Grooming* e Reconfiguração de Rede

Em conformidade com os objetivos deste trabalho (ver seção 1.2), esta seção apresenta uma extensão dos modelos propostos anteriormente para um horizonte de planejamento limitado.

Conforme destacado na revisão da literatura (ver subseção 2.3.2), no **TGP** dinâmico em que se leva em consideração questões sobre a reconfiguração da rede, utiliza-se um conjunto de matrizes para caracterizar o tráfego. O padrão de tráfego pode mudar dentro deste conjunto de matrizes durante um certo período de tempo (por exemplo, ao longo de um dia, uma semana ou um mês), tornando-se necessário realizar uma eventual reconfiguração da rede quando o padrão de tráfego se altera de uma matriz para outra do conjunto. Por reconfiguração entende-se uma modificação da topologia virtual para atendimento das novas requisições realizada preferencialmente sem a interrupção das conexões já existentes (por exemplo, através da utilização dos comprimentos de onda já em uso ou da alocação de novos comprimentos de onda).

Este problema é denominado de *Traffic Grooming and Reconfiguration Problem* – **TGRP** e sua abordagem por completo foge do escopo deste trabalho. Contudo, uma investigação preliminar foi conduzida procurando-se determinar a adequação das abordagens propostas para os modelos anteriores em resolver o **TGRP**.

Em particular, apesar do padrão de tráfego poder variar de diferentes maneiras, neste trabalho só foi considerado um padrão de *tráfego incremental*, conforme definição adotada em [55]. Entende-se por padrão de *tráfego incremental* aquele que é dinâmico e que, porém, nunca é terminado. Isto representa, segundo os autores em [55], a situação em que se espera que os tráfegos tenham uma longa duração (*holding time*) como no caso do provisionamento de conexões de alta velocidade.

Para se formular matematicamente o **TGRP** deve-se antes construir uma representação estendida da rede conforme foi discutido na subseção 3.1.1. A subseção 3.4.1 introduz a notação utilizada, bem como uma formulação de programação matemática para o problema. Semelhante ao que foi realizado antes, tal formulação pode ser simplificada pela adoção de uma representação em camadas (conforme discutido na subseção 3.1.3), resultando em uma formulação condensada do problema (subseção 3.4.2) que é utilizada no restante deste trabalho.

3.4.1 Formulação Matemática do TGRP

Considere a seguinte notação a ser utilizada na formulação de programação matemática associada ao **TGRP**, cuja representação estendida seja $G = (N, A)$:

- N representa o conjunto de nós, que pode ser particionado em dois subconjuntos disjuntos, $N = N^e \cup N^o$, em que N^e representa o conjunto de nós de *add-drop* e N^o representa o conjunto de nós óticos;
- A representa o conjunto de arcos, que pode ser particionado em dois, $A = A^e \cup A^o$, em que A^e representa o conjunto de arcos entre nós de *add-drop* e os nós óticos, enquanto que A^o representa o conjunto de arcos entre nós óticos;
- T representa o conjunto de períodos, tal que $T = \{1, 2, \dots, n_t\}$, em que n_t representa o número total de períodos;
- C representa o conjunto de *containers virtuais* (ou comprimentos de onda), tal que $C = \{1, 2, \dots, W\}$, em que W representa o número total de *containers virtuais* (ou comprimentos de onda) disponíveis;
- P representa o conjunto de produtos a serem transportados, tal que $P = \{t_p = (s_p, d_p, f_p^1, f_p^2, \dots, f_p^{n_t}, m_p) \mid s_p \in N^e, d_p \in N^e, m_p \in \mathbb{Z}_+, 1 \leq p \leq n_p \text{ e } f_p^t \in \mathbb{Z}_+, 1 \leq p \leq n_p, 1 \leq t \leq n_t\}$, em que n_p é o número total de produtos e $f_p^{t-1} \leq f_p^t, t = 2, \dots, n_t, \forall t_p \in P$;
- s_p representa o nó de origem do produto $t_p \in P$;
- d_p representa o nó de destino do produto $t_p \in P$;
- f_p^t representa a quantidade (em número de canais) do produto $t_p \in P$ que deve ser transportada de s_p a d_p no período $t \in T$;
- m_p representa a quantidade máxima (em número de canais) do produto $t_p \in P$ que pode ser transportada em um *container virtual* (ou comprimento de onda);
- c_{ijk}^t representa o custo de utilização do *container virtual* (ou comprimento de onda) $k \in C$ no arco $(i, j) \in A$ no período $t \in T$.

Cumprir destacar que a suposição do tráfego ser incremental está representada na definição de P , em que o tráfego de um período não é menor que o tráfego do período anterior, ou ainda, $f_p^{t-1} \leq f_p^t, t = 2, \dots, n_t, \forall t_p \in P$. Além disso, as seguintes variáveis são utilizadas na formulação do **TGRP**:

- f_{ijk}^{pt} representa a quantidade (em canais) do produto $t_p \in P$ que trafega através do arco $(i, j) \in A$ utilizando o *container virtual* $k \in C$ no período $t \in T$;
- w_{ijk}^t indica o uso do *container virtual* $k \in C$ no arco $(i, j) \in A$ no transporte de algum produto no período $t \in T$.

Por fim, considere que a capacidade de um *container virtual* (ou comprimento de onda) será dada por:

$$\Delta = \text{mmc}^3 \{ m_p \mid \forall t_p = (s_p, d_p, f_p^1, f_p^2, \dots, f_p^{n_t}, m_p) \in P \}.$$

Ao passo que a contribuição (ou melhor, a ocupação) correspondente a uma unidade de fluxo de um produto transportado em um *container virtual* é representada

³Mínimo múltiplo comum.

por:

$$\delta_p = \Delta/m_p, \quad \forall t_p \in P.$$

Sendo assim, a formulação de programação matemática, $TGRP$, associada ao problema é dada por:

$$(TGRP) \quad \min \sum_{t \in T} \sum_{k \in C} \sum_{(i,j) \in A} c_{ijk}^t w_{ijk}^t \quad (3.18a)$$

sujeito a:

$$\sum_{k \in C} \sum_{(i,j) \in A_e} f_{ijk}^{pt} - \sum_{k \in C} \sum_{(j,i) \in A_e} f_{jik}^{pt} = \begin{cases} f_p^t, i = s_p \\ f_p^t, i = d_p \\ 0, i \neq s_p \neq d_p \end{cases}, \forall t_p \in P, \forall i \in N^e, \forall t \in T \quad (3.18b)$$

$$\sum_{(i,j) \in A} f_{ijk}^{pt} - \sum_{(j,i) \in A} f_{jik}^{pt} = 0, \forall k \in C, \forall t_p \in P, \forall i \in N^o, \forall t \in T \quad (3.18c)$$

$$\sum_{t_p \in P} \delta_p f_{ijk}^{pt} \leq \Delta \sum_{t=1}^t w_{ijk}^t, \forall k \in C, \forall (i,j) \in A, \forall t \in T \quad (3.18d)$$

$$\sum_{(i,j) \in A} w_{ijk}^t - \sum_{(j,i) \in A} w_{jik}^t = 0, \forall k \in C, \forall i \in N^o, \forall t \in T \quad (3.18e)$$

$$\sum_{t \in T} w_{ijk}^t \leq 1, \forall k \in C, \forall (i,j) \in A \quad (3.18f)$$

$$f_{ijk}^{pt} \geq 0, \forall k \in C, \forall t_p \in P, \forall (i,j) \in A, \forall t \in T \quad (3.18g)$$

$$w_{ijk}^t \in \{0, 1\}, \forall k \in C, \forall (i,j) \in A, \forall t \in T \quad (3.18h)$$

$$f_{ijk}^{pt} \text{ inteiro}, \forall k \in C, \forall t_p \in P, \forall (i,j) \in A, \forall t \in T \quad (3.18i)$$

A função objetivo dada por (3.18a) procura minimizar o custo total de utilização e/ou alocação dos *containers virtuais*. As restrições (3.18b) e (3.18c) garantem a conservação de fluxo dos produtos para os nós de *add-drop* e para os nós óticos. Vale ressaltar que nas restrições (3.18b) todo o fluxo que entra ou sai de um nó de *add-drop* (independentemente do comprimento de onda utilizado para transportá-lo) é considerado nos somatórios, permitindo assim que o fluxo de um produto, que entre em um nó de *add-drop* utilizando um dado comprimento de onda, venha a deixar o mesmo nó através de um outro comprimento de onda. Já o mesmo não se aplica aos nós óticos, uma vez que as restrições (3.18c) são escritas separadamente para cada comprimento de onda $k \in C$.

As restrições (3.18d) impõem um limite sobre o volume total de produtos transportados através de um *container virtual* no arco $(i,j) \in A$, isto é, um limite sobre a capacidade dos *containers virtuais*, além de estabelecer uma ligação entre as variáveis de fluxo (f_{ijk}^{pt}) e de decisão (w_{ijk}^t). Já as restrições (3.18e) garantem a continuidade dos *lightpaths*, além de serem responsáveis, juntamente com a estrutura do grafo, pela não bifurcação dos mesmos. As restrições (3.18f) garantem que a alocação de um *container virtual* ocorra em um único período.

Por trás da utilização das restrições (3.18e) e (3.18f) está a opção de se considerar que uma vez alocada uma capacidade em um dado período (ou ainda, alocado um *container virtual*) não poderá ocorrer outra alocação no futuro e a capacidade

(*container virtual*) estará disponível em todos os períodos seguintes ao da alocação inicial. Vale dizer também que o custo de alocação só será contabilizado para o período inicial de alocação.

Finalmente, as restrições (3.18g), (3.18h) e (3.18i) definem as variáveis de fluxo (f_{ijk}^{pt}) como variáveis inteiras e não negativas, enquanto que as variáveis de decisão (w_{ijk}^t) são definidas como binárias.

Como mencionado acima, uma vez alocada uma capacidade (alocado um *container virtual*), ela estará disponível para todos os demais períodos no futuro. Isto contribui em parte para que o roteamento de períodos anteriores não seja drasticamente modificado apesar do modelo, a princípio, não fornecer nenhuma outra garantia explícita disso.

Já em relação a função de custo adotada, deve-se destacar que ela é semelhante a discutida anteriormente; contudo, é comum se supor a utilização de uma taxa de desconto que faz com que o custo de alocação nos primeiros períodos seja maior que nos últimos. Isso, por sua vez, contribui para que a alocação seja adiada o máximo possível, uma vez que será mais barato realizá-la no futuro.

3.4.2 Formulação Condensada do TGRP

A adoção da representação em camadas (conforme discutido na subseção 3.1.3) permite reformular o problema de forma mais simples e compacta. Para tanto, deve-se considerar a seguinte notação a ser utilizada na formulação matemática condensada para o **TGRP** cuja representação em camadas, conforme descrito na subseção 3.1.3, seja dada pelo grafo $G_\ell = (N_\ell, A_\ell)$:

- N_ℓ representa o conjunto de nós da representação em camadas, que pode ser particionado em dois subconjuntos disjuntos, $N = N_\ell^e \cup N_\ell^o$, em que N_ℓ^e representa o conjunto de nós de *add-drop* e N_ℓ^o representa o conjunto de nós óticos independentemente das camadas;
- A_ℓ representa o conjunto de arcos da representação em camadas;
- c_ℓ^t representa o custo de utilização do arco $\ell \in A_\ell$ no período $t \in T$.

As tuplas do conjunto de produtos devem ser ajustadas de modo que os nós de origem e destino de produtos correspondam aos nós pertencentes a N_ℓ^e .

Além disso, as seguintes variáveis são utilizadas na formulação condensada do **TGRP**:

- f_ℓ^{pt} representa a quantidade (em canais) do produto $t_p \in P$ que trafega através do arco $\ell \in A_\ell$ no período $t \in T$;
- w_ℓ^t indica o uso do arco $\ell \in A_\ell$ no transporte de algum produto no período $t \in T$.

Como antes, a capacidade máxima de qualquer um dos arcos é dada por Δ (a mesma capacidade máxima atribuída antes a um *container virtual*) e cada unidade de fluxo de um produto continua a consumir δ_p unidades de capacidade de transporte.

Por fim, novamente, $A_\ell^+(i)$ é utilizado para representar o conjunto de todos os arcos que saem de um nó $i \in N_\ell$, enquanto que $A_\ell^-(i)$ representa o conjunto de todos os arcos que chegam em um nó $i \in N_\ell$, isto é:

$$A_\ell^+(i) = \{\ell \in A_\ell \mid \ell = (i, j), j \in N_\ell\}, \forall i \in N_\ell,$$

$$A_\ell^-(i) = \{\ell \in A_\ell \mid \ell = (j, i), j \in N_\ell\}, \forall i \in N_\ell.$$

Dessa forma, uma formulação condensada, $TGRP_C$, associada ao problema é dada por:

$$(TGRP_C) \quad \min \sum_{t \in T} \sum_{\ell \in A_\ell} c_\ell^t w_\ell^t \quad (3.19a)$$

sujeito a:

$$\sum_{\ell \in A_\ell^+(i)} f_\ell^{pt} - \sum_{\ell \in A_\ell^-(i)} f_\ell^{pt} = b_i^{pt}, \quad \forall t_p \in P, \forall i \in N_\ell, \forall t \in T \quad (3.19b)$$

$$\sum_{\ell \in A_\ell^+(i)} w_\ell^t - \sum_{\ell \in A_\ell^-(i)} w_\ell^t = 0, \quad \forall i \in N_\ell^o, \forall t \in T \quad (3.19c)$$

$$\sum_{t_p \in P} \delta_p f_\ell^{pt} \leq \Delta \sum_{t=1}^t w_\ell^t, \quad \forall \ell \in A_\ell, \forall t \in T \quad (3.19d)$$

$$\sum_{t \in T} w_\ell^t \leq 1, \quad \forall \ell \in A_\ell \quad (3.19e)$$

$$f_\ell^{pt} \geq 0, \quad \forall t_p \in P, \forall \ell \in A_\ell, \forall t \in T \quad (3.19f)$$

$$w_\ell^t \in \{0, 1\}, \quad \forall \ell \in A_\ell, \forall t \in T \quad (3.19g)$$

$$f_\ell^{pt} \text{ inteiro}, \quad \forall t_p \in P, \forall \ell \in A_\ell, \forall t \in T \quad (3.19h)$$

em que b_i^{pt} é dado por:

$$b_i^{pt} = \begin{cases} f_p^t, & \text{se } i = s_p \\ -f_p^t, & \text{se } i = d_p \\ 0, & \text{se } i \neq s_p \neq d_p \end{cases} \quad \forall t_p \in P, \forall i \in N_\ell, \forall t \in T. \quad (3.20)$$

Cumpra reafirmar que a formulação $TGRP_C$ dada por (3.19a)–(3.19h) é totalmente equivalente àquela apresentada anteriormente, $TGRP$, através das equações (3.18a)–(3.18i).

A função objetivo dada por (3.19a) procura minimizar o custo total de utilização/alocação dos arcos da representação em camadas. As restrições (3.19b) garantem a conservação de fluxo dos produtos para todos os nós da representação em camadas. Já as restrições (3.19c) garantem que a quantidade de arcos alocados que saem de um nó ótico é igual a quantidade de arcos alocados que chegam no mesmo nó (sendo, assim, semelhantes às restrições (3.18e), responsáveis pela continuidade dos *lightpaths*). As restrições (3.19d), por sua vez, impõem um limite sobre o volume total de produtos transportados através de um arco $\ell \in A_\ell$, além de estabelecer uma ligação entre as variáveis de fluxo (f_ℓ^{pt}) e de decisão (w_ℓ^t), enquanto as restrições (3.19e) limitam a alocação do container virtual a no máximo um período.

Finalmente, as restrições (3.19f), (3.19g) e (3.19h) definem as variáveis de fluxo (f_ℓ^{pt}) como variáveis inteiras e não negativas, enquanto que as variáveis de decisão (w_ℓ^t) são definidas como binárias.

3.4.3 Limitando o Número de GOXCs

De forma análoga ao que foi realizado na seção 3.2 para o **TGP**, pode-se também limitar a realização de *grooming* a um subconjunto dos elementos da rede no **TGRP**.

Dessa forma, adotando-se um procedimento similar ao que foi realizado na seção 3.2, as formulações *STGRP* e *STGRP_C* são obtidas. Essas formulações representam, respectivamente, as formulações “natural” e condensada do **TGRP** em que a capacidade de realização de *grooming* foi limitada a um subconjunto dos nós. Contudo, tais formulações não são apresentadas neste texto, uma vez que representam extensões diretas das formulações de **TGRP**, *TGRP* e *TGRP_C*, aplicando-se as idéias desenvolvidas na seção 3.2.

3.4.4 Limitando o Número de *Transceivers*

Por outro lado, à semelhança do que foi realizado na seção 3.3, pode-se ajustar as formulações propostas para **TGRP** de modo a se representar as situações em que a capacidade de realização de *grooming* é limitada pelo número de *transceivers* (ou ADMs) existentes em cada elemento da rede.

Sendo assim, para se impor uma limitação sobre número de *transceivers* existentes em cada elemento da rede, deve-se acrescentar o seguinte conjunto de restrições:

$$\sum_{k \in C} \sum_{(i,j) \in A^e} w_{ijk}^t \leq N_i^T, \forall i \in N^e, \forall t \in T; \quad (3.21)$$

$$\sum_{k \in C} \sum_{(j,i) \in A^e} w_{jik}^t \leq N_i^R, \forall i \in N^e, \forall t \in T. \quad (3.22)$$

em que N_i^T e N_i^R representam o número de transmissores e receptores óticos, respectivamente, existentes no elemento da rede associado ao nó de *add-drop* $i \in N^e$.

Tais restrições podem ser utilizadas tanto em *TGRP* como em *STGRP*; porém, para utilizá-las em conjunto com as formulações condensadas, *TGRP_C* e *STGRP_C*, torna-se necessário reescrevê-las da seguinte forma:

$$\sum_{\ell \in A_\ell^+(i)} w_\ell^t \leq N_i^T, \forall i \in N_\ell^e, \forall t \in T; \quad (3.23)$$

$$\sum_{\ell \in A_\ell^-(i)} w_\ell^t \leq N_i^R, \forall i \in N_\ell^e, \forall t \in T. \quad (3.24)$$

3.5 Considerações Finais

Neste capítulo procurou-se desenvolver e apresentar uma formulação de programação matemática para o **TGP**. Tal formulação “natural” foi simplificada pela adoção de uma representação em camadas produzindo uma formulação condensada bem mais próxima de vários problemas discutidos na literatura nos últimos anos.

Essa formulação condensada é utilizada no restante deste trabalho visto que ela é totalmente equivalente à formulação original.

Além disso, formulações para variações do problema original, em especial para **STGP** e para **TGRP**, também foram apresentadas e discutidas. Cumpre dizer que,

de forma análoga ao **TGP**, o restante deste trabalho utiliza as versões condensadas das formulações propostas para esses problemas.

O próximo capítulo discute as duas abordagens a serem adotadas na resolução do **TGP**, procurando apresentar e analisar os principais aspectos de cada uma delas.

Capítulo 4

Abordagens Propostas

COMO foi exposto anteriormente, o objetivo deste trabalho é desenvolver algoritmos de resolução para o **TGP**, além de testar o desempenho (qualidade dos resultados obtidos) dos mesmos.

Para tanto, neste trabalho são adotadas duas abordagens formais distintas baseadas em técnicas bastante consolidadas na atualidade. Na primeira delas faz-se uso da relaxação lagrangeana para obtenção de limites inferiores a serem utilizados, por sua vez, em um procedimento de busca (algoritmo de *branch-and-bound*). Já na segunda, uma análise da estrutura facial do poliedro associado ao conjunto de soluções do problema foi realizada de modo a identificar inequações válidas e, eventualmente, facetas (cortes faciais) que, por sua vez são utilizadas em um outro procedimento de busca (algoritmo de *branch-and-cut*).

Cumprе destacar que tais abordagens guardam entre si uma certa relação uma vez que um algoritmo de *branch-and-cut* pode ser visto como uma combinação de um método de *planos-de-corte* com um algoritmo de *branch-and-bound*. A principal diferença funcional entre eles se encontra na forma de obter (e atualizar) os limites inferiores utilizados na busca. O uso da relaxação lagrangeana permite uma decomposição do problema original em subproblemas mais simples cujas soluções são, então, utilizadas na obtenção de um limite inferior para o problema original. Já na abordagem poliédrica uma relaxação do problema original (por exemplo, a relaxação da integralidade das variáveis – relaxação linear) é usada como base para obtenção de uma aproximação inicial do fecho inteiro do poliedro associado ao conjunto de soluções do problema. Essa aproximação é melhorada através do acréscimo de novos *planos-de-corte* (inequações válidas que eliminam ou “cortam” soluções indesejáveis do espaço de busca).

Na abordagem lagrangeana, a simplicidade e, até mesmo, facilidade teórica com que são obtidos e calculados os limites inferiores acaba por se refletir em uma maior complexidade dos métodos de atualização (maximização) dos mesmos, principalmente no caso de se desejar um sequência de limites inferiores monotonicamente não decrescente. Já na abordagem poliédrica, um maior esforço (dada a complexidade teórica) no sentido de se obter “bons” *planos-de-corte* traz consigo a vantagem de produzir naturalmente tal sequência de limites não decrescentes.

Nas duas abordagens, a geração de soluções viáveis é feita através de heurísticas sendo que tais soluções são utilizadas para controlar e, até mesmo, interromper o processo de busca. Preferencialmente essas heurísticas devem lançar mão de informações dos limites inferiores — gerados pela resolução dos subproblemas, na relaxação lagrangeana, e do problema relaxado, na abordagem poliédrica — para obter soluções viáveis de boa qualidade (próximas da solução ótima).

A seguir, passa-se a relatar as diversas opções adotadas em cada uma das abordagens usadas na resolução do **TGP** neste trabalho.

4.1 Algoritmo de *Branch-and-Bound*

O desenvolvimento e aplicação de um algoritmo de *branch-and-bound* baseado no uso da relaxação lagrangeana é motivado por duas razões distintas. Em primeiro lugar está o fato dessa abordagem ter sido aplicada com sucesso a inúmeros problemas de otimização combinatória ao longo das últimas décadas e, em especial, a diversos problemas relacionados ao planejamento e projeto de redes [74]. Uma outra razão para sua adoção reside na facilidade de se obter uma implementação funcional, desde que não se deseje uma sequência de limites inferiores monotonicamente não decrescente. Isto se deve, pelo menos em parte, ao fato de que o problema original é decomposto em subproblemas mais simples¹.

A técnica de *branch-and-bound* (ou de enumeração implícita) foi introduzida para solução de problemas de programação inteira mista em [39, 93] e refinada nos anos seguintes (ver, por exemplo, [52]). Consiste em uma abordagem de *divisão-e-conquista* que procura resolver o problema original através de sua subdivisão em problemas menores para os quais limites inferiores e superiores são calculados. Uma parte crucial do sucesso de um algoritmo de *branch-and-bound* reside no cálculo dos limites inferiores (no caso de um problema de minimização). Dessa forma, o conceito de relaxação se torna fundamental. A partir deste ponto, toda a discussão é realizada objetivando problemas de minimização, tendo em vista as formulações apresentadas no capítulo anterior. Contudo, extensões para problemas de maximização podem ser obtidas facilmente.

Definição 4.1. *Dado um conjunto finito E , um conjunto \mathcal{I} de subconjuntos de E (soluções viáveis) e uma função $c : E \rightarrow \mathbb{R}$. Para cada conjunto $F \subseteq E$ seja $c(F) = \sum_{e \in F} c(e)$. Então o problema de otimização combinatória linear (POCL) consiste em encontrar um conjunto $I^* \in \mathcal{I}$ tal que*

$$c(I^*) = \min\{c(I) \mid I \in \mathcal{I}\}.$$

Pode-se representar esse POCL pela tripla (E, \mathcal{I}, c) .

Definição 4.2. *Considere que sejam dados dois problemas de otimização combinatória lineares (E, \mathcal{I}, c) e (E', \mathcal{I}', c') e uma função injetora $\varphi : E \rightarrow E'$. (\mathcal{I}', c') é uma relaxação de (\mathcal{I}, c) se $\varphi(I) \in \mathcal{I}'$ e $c(I) = c'(\varphi(I))$ para todo $I \in \mathcal{I}$.*

¹Apesar disso, a obtenção de uma implementação competitiva pode exigir um certo investimento adicional, quer seja na codificação dos algoritmos, quer seja na definição das estruturas de dados.

Dessa forma, a solução do problema relaxado fornece um limite inferior (no caso de um problema de minimização) para o valor ótimo da função objetivo do problema original. Quanto mais “justa” for a relaxação, melhor será esse limite. Contudo, uma relaxação só é útil quando for tratada eficientemente por algum algoritmo. A partir da remoção da restrição de integralidade sobre as variáveis de uma formulação de programação inteira de um POCL, obtém-se uma relaxação de programação linear.

Antes de se detalhar o funcionamento de um algoritmo de *branch-and-bound*, deve-se estabelecer algumas terminologias sobre os limites inferiores (gerados a partir da resolução de relaxações) e superiores (obtidos através da busca de soluções viáveis). Um limite é dito **local** se ele for válido somente para um subproblema, enquanto que ele será **global** caso seja um limite válido para o problema original. A partir da resolução de uma relaxação do problema corrente obtém-se um limite inferior local para o valor da função objetivo. Se a solução obtida para a relaxação for viável para o problema original (ou puder ser transformada em uma) que possua um valor para a função objetivo menor do que qualquer outra solução viável já encontrada, ela deve ser armazenada e o limite superior global para o valor da função objetivo deve ser atualizado.

Um algoritmo de *branch-and-bound* mantém uma lista de subproblemas (gerados a partir do problema original), que inicialmente contém o próprio problema original. A cada iteração o algoritmo seleciona um subproblema dessa lista, calcula um limite inferior local para esse subproblema e procura melhorar o limite superior global. Caso o limite inferior local seja maior que o limite superior global o subproblema pode ser abandonado, uma vez que sua solução não poderá ser melhor que a melhor solução viável encontrada até o presente momento. Caso contrário, deve-se verificar se a solução da relaxação do subproblema é uma solução viável para o problema original. Neste caso, o subproblema foi completamente resolvido e, portanto, também pode ser abandonado. Se o limite inferior local não exceder o limite superior global e nenhuma solução encontrada para o subproblema corrente for viável para o problema original, uma ramificação (*branching*) do subproblema deve ser realizada. Essa ramificação é feita através da divisão do subproblema atual em uma coleção de novos subproblemas cuja união de seus conjuntos de soluções viáveis contém todas as soluções viáveis do subproblema corrente. No momento em que a lista de subproblemas se tornar vazia, a solução viável armazenada e cujo valor corresponde ao limite superior global será a solução ótima do problema original.

Pelo que foi exposto, pode-se concluir que para a eficiência de um algoritmo de *branch-and-bound* é importante não só a qualidade da relaxação utilizada, mas também a qualidade das soluções viáveis geradas durante a busca, uma vez que elas podem auxiliar no processo de busca limitando o número de subproblemas a serem explorados. Nas subseções seguintes são fornecidos os detalhes sobre a implementação do algoritmo de *branch-and-bound* baseado no uso da relaxação lagrangeana.

4.1.1 Limites Inferiores

Nesta primeira abordagem os limites inferiores são obtidos através do uso da relaxação lagrangeana. Esta subseção descreve brevemente a utilização de tal técnica na obtenção de limites inferiores para problemas de programação linear inteira, maiores

detalhes podem ser encontrados em [115].

Nesta técnica, a partir da introdução de multiplicadores de Lagrange, algumas restrições do problema original são relaxadas (removidas da formulação) e acrescentadas à função objetivo (utilizando-se os respectivos multiplicadores de Lagrange como pesos). Considere o seguinte problema P de programação linear inteira:

$$Z = \min \{c^T x : Ax \leq b, x \in \mathbb{Z}_+^n\}$$

em que c e $b \in \mathbb{R}^n$ e A é uma matriz $m \times n$.

Suponha que o conjunto de restrições desse problema possa ser particionado em dois subconjuntos de restrições $A_1x \leq b_1$ e $A_2x \leq b_2$, em que a matriz A_1 possui $m_1 < m$ linhas. Além disso, suponha que a seguinte relaxação

$$\min \{c^T x : A_2x \leq b_2, x \in \mathbb{Z}_+^n\}$$

obtida a partir da remoção das restrições $A_1x \leq b_1$ é facilmente resolvida.

Sendo assim, dado um vetor de multiplicadores de Lagrange $\lambda \in \mathbb{R}_+^{m_1}$, pode-se definir a relaxação lagrangeana do problema P em relação às restrições $A_1x \leq b_1$ da seguinte forma:

$$L(\lambda) = \min \{c^T x + \lambda^T (A_1x - b_1) : A_2x \leq b_2, x \in \mathbb{Z}_+^n\}.$$

Vale notar que $\lambda^T (A_1x - b_1) \leq 0$, para todo x viável no problema original, uma vez que $\lambda \geq 0$. Logo $L(\lambda) \leq Z$ para todo x viável no problema original, isto é, $L(\lambda)$ representa um limite inferior para Z .

Além disso, $L(0)$ equivale ao problema obtido removendo-se as restrições $A_1x \leq b_1$; entretanto, um limite inferior maior pode ser obtido a partir da escolha de um $\lambda \neq 0$. A obtenção do valor de λ que resulta no maior limite inferior possível, dá origem ao problema lagrangeano dual, P^D , que pode ser representado da seguinte forma:

$$L^D = \max \{L(\lambda) : \lambda \geq 0\}.$$

É possível demonstrar que o problema lagrangeano dual corresponde a um problema de maximização de uma função côncava linear por partes e que sua solução, L^D , também é um limite inferior para o problema original, ou ainda, $L^D \leq Z$ [115]. Infelizmente, a função objetivo do problema lagrangeano dual é não-diferenciável, o que dificulta a resolução do mesmo, como se verá mais adiante.

Maiores detalhes sobre o uso da relaxação lagrangeana, bem como características da função lagrangeana, do problema lagrangeano dual e de suas relações com o problema original podem ser encontradas em [53, 115].

No caso específico das formulações propostas no capítulo anterior foram implementados e avaliados quatro limites inferiores para o **TGP** obtidos a partir da relaxação lagrangeana de subconjuntos distintos das restrições. Dessa maneira, foi possível se averiguar e comparar adequadamente a qualidade dos limites obtidos.

4.1.2 Limites Superiores

Como visto anteriormente, a eficiência do processo de busca também depende diretamente da qualidade dos limites superiores. Sendo assim, para se obter limites

superiores de “boa qualidade”, um conjunto de heurísticas simples e eficientes foi desenvolvido e testado.

Tais heurísticas utilizam informações obtidas durante o cálculo de limites inferiores (resolução dos problemas relaxados) procurando, assim, considerar o andamento do processo de busca na geração de limites superiores. Contudo, como as soluções dos problemas relaxados não correspondem necessariamente a soluções viáveis para o problema original, um primeiro esforço é, então, direcionado na viabilização de tais soluções.

Uma vez que se consiga viabilizar heurísticamente uma solução gerada a partir do cálculo de um limite inferior, outras heurísticas foram aplicadas visando melhorar a qualidade da solução encontrada (isto é, reduzir seu custo).

4.1.3 Problema Lagrangeano Dual

Como visto na subseção 4.1.1, para se obter limites inferiores de qualidade, torna-se necessário resolver o problema lagrangeano dual P^D . Todavia o fato deste problema apresentar uma função objetivo não-diferenciável torna essa tarefa mais difícil pois, apesar desse problema não apresentar restrições além da não negatividade dos multiplicadores de Lagrange, a ausência de derivadas (para determinados pontos de L) impede a aplicação direta do ferramental padrão oriundo do cálculo diferencial. Neste momento, torna-se necessário a utilização de um ferramental oriundo da análise convexa de modo a se contornar as dificuldades impostas pela ausência de uma direção de subida (gradiente).

Para resolver o problema lagrangeano dual P^D , qualquer método de otimização não-diferenciável utiliza-se de informações geradas por um “oráculo”. Mais especificamente, para um dado vetor de multiplicadores de Lagrange $\bar{\lambda}$, o “oráculo” é capaz de calcular para uma função convexa f (neste caso, $f \equiv -L$) tanto o seu valor quanto o de um de seus *subgradientes* no ponto $\bar{\lambda}$. Desse modo, é conveniente introduzir a noção de *subdiferencial* (e de *subgradientes*).

Definição 4.3. *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa e $\bar{\lambda} \in \mathbb{R}^n$. Então $\xi \in \mathbb{R}^n$ é um **subgradiente** de f no ponto $\bar{\lambda}$ se*

$$f(\lambda) \geq f(\bar{\lambda}) + \xi^T(\lambda - \bar{\lambda}), \forall \lambda \in \mathbb{R}^n.$$

*Além disso, pode-se definir a **subdiferencial** de f no ponto $\bar{\lambda}$, ou ainda, $\partial f(\bar{\lambda})$ da seguinte forma:*

$$\partial f(\bar{\lambda}) = \{ \xi \in \mathbb{R}^n \mid \xi^T(\lambda - \bar{\lambda}) \leq f(\lambda) - f(\bar{\lambda}), \forall \lambda \in \mathbb{R}^n \}.$$

Infelizmente, muitos dos métodos de otimização não-diferencial baseados no uso de subgradientes apresentam, por vezes, sérios problemas de convergência [126]. Esse fenômeno parece estar associado ao fato de que, durante a busca da solução ótima, tais métodos procuram reduzir a distância entre o ponto atual e o ponto de ótimo da função, sem contudo garantir um decréscimo de valor da mesma [126]. Para se garantir que a cada passo da busca ocorra um decréscimo do valor da função objetivo torna-se necessário introduzir o conceito de ε -*subdiferencial* (e de ε -*subgradientes*) cujas propriedades permitem contornar as dificuldades impostas pela não-diferenciabilidade da função objetivo (ver [126]).

Definição 4.4. Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa e $\bar{\lambda} \in \mathbb{R}^n$. Então, para um dado $\varepsilon \geq 0$, $\xi \in \mathbb{R}^n$ é um ε -**subgradiente** de f no ponto $\bar{\lambda}$ se

$$f(\lambda) \geq f(\bar{\lambda}) + \xi^T(\lambda - \bar{\lambda}) - \varepsilon, \forall \lambda \in \mathbb{R}^n.$$

Além disso, pode-se definir a ε -**subdiferencial** de f no ponto $\bar{\lambda}$, ou ainda, $\partial_\varepsilon f(\bar{\lambda})$ da seguinte forma:

$$\partial_\varepsilon f(\bar{\lambda}) = \{\xi \in \mathbb{R}^n \mid \xi^T(\lambda - \bar{\lambda}) - \varepsilon \leq f(\lambda) - f(\bar{\lambda}), \forall \lambda \in \mathbb{R}^n\}.$$

Existem várias abordagens de otimização não-diferenciável (usando *subgradientes* ou ε -*subgradientes*) que podem ser aplicadas para se solucionar o problema lagrangeano dual P^D , tais como: método de subgradiente (ou método de Shor) [141], método de *plano de corte* (ou secantes) [29, 89], método de centros analíticos [57, 97] e método de feixe (ou *bundle*) [71]. Cada um desses métodos possui vantagens e desvantagens e sua eficiência parece estar relacionada a natureza do problema a ser resolvido. Pode-se encontrar em [96] uma descrição concisa de cada um deles. Em particular, os métodos de subgradiente são notórios não só por sua simplicidade (se comparado aos demais), mas também pela ausência de critérios de parada “bem definidos”. Por outro lado, os métodos de feixe são considerados mais robustos e precisos; entretanto, cada iteração é muito mais complexa (se comparado ao anterior), uma vez que a direção de busca é determinada através da resolução de um problema de programação quadrática. Neste trabalho, optou-se pela utilização do método de subgradiente (também chamado de método de Shor ou Kiev), uma vez que os experimentos computacionais demonstraram que tal abordagem apresentava uma “boa” convergência quando aplicada à resolução do **TGP**.

4.1.4 Heurística Lagrangeana

Um dos primeiros resultados deste trabalho foi a implementação e teste de uma heurística lagrangeana, em que o **TGP** foi relaxado (via relaxação lagrangeana) para o cálculo de limites inferiores e tais limites foram atualizados através de um método de resolução do problema lagrangeano dual (através do método de subgradiente). Além disso, soluções viáveis foram obtidas através de uma heurística em que informações sobre os limites inferiores são utilizadas (conforme descrito anteriormente).

Tal heurística lagrangeana pode ser vista como uma restrição da implementação de um algoritmo de *branch-and-bound*, em que apenas o nó inicial (correspondendo ao problema original) é explorado. Apesar de tal abordagem não garantir a obtenção de uma solução ótima, os resultados dos testes com essa heurística são muito promissores e são apresentados em detalhe mais adiante. De todo modo, para os casos em que essa heurística falhar (isto é, quando uma solução ótima não for encontrada), torna-se necessário a adoção de um esquema completo de busca no espaço de soluções.

4.1.5 Algoritmo de Busca

Por fim, o objetivo maior desta primeira abordagem foi o desenvolvimento e teste de um algoritmo de *branch-and-bound* baseado no uso da relaxação lagrangeana para resolução do **TGP**.

Contudo, experimentos iniciais demonstraram que sua eficiência podia ser comprometida. Isto se devia em parte à ausência de critérios bem estabelecidos para o caminhamento na árvore de subproblemas. Além disso, mecanismos de redução do espaço de soluções também pareciam ser necessários para uma melhoria de desempenho do processo de busca.

Sendo assim, tornou-se imprescindível o estudo e desenvolvimento de mecanismos de redução do espaço de busca (como, por exemplo, a “fixação” de variáveis), bem como o estabelecimento de estratégias eficazes para o caminhamento na árvore de subproblemas, para a obtenção de um procedimento de busca eficiente.

4.2 Algoritmo de *Branch-and-Cut*

A segunda abordagem desenvolvida neste trabalho, um algoritmo de *branch-and-cut*, como mencionado anteriormente, apresenta similaridades em relação à primeira pelo fato de ambas se utilizarem de um procedimento de busca sistemática no espaço de soluções. Contudo, nesta abordagem o cálculo de limites inferiores é realizado através de sucessivas aproximações do fecho inteiro do poliedro associado ao conjunto de soluções viáveis. Antes de se descrever o desenvolvimento e implementação de um algoritmo de *branch-and-cut* para o **TGP**, apresentar-se-á um breve histórico da aplicação de tais técnicas na resolução de problemas de otimização combinatória.

Primeiramente, é preciso ressaltar a possibilidade de transformação de certos problemas de otimização combinatória em problemas de programação linear sobre um poliedro inteiro (mesmo que definido implicitamente). De modo a se solucionar tais problemas de programação linear torna-se necessário obter uma descrição do poliedro associado ao conjunto de soluções viáveis. Sabe-se, ainda, que uma descrição não-redundante de um poliedro é dada por um sistema de inequações que definem facetas do mesmo. Porém, para problemas \mathcal{NP} -difíceis é pouco provável que se consiga obter uma descrição completa deste tipo [87, 118], uma vez que a obtenção de uma descrição completa de um politopo associado a um problema \mathcal{NP} -difícil juntamente com um algoritmo polinomial para resolução de seu problema de separação implicaria em $\mathcal{P} = \mathcal{NP}$.

Deve-se dizer que tal estratégia de se associar um politopo a um problema, e mais especificamente, encontrar um sistema de inequações que descreva tal poliedro (definido implicitamente pelo seus vértices) constitui um tópico central da pesquisa em combinatória poliédrica [48]. Ainda segundo [48], o desenvolvimento desta área de investigação pode ser dividido em 3 períodos.

O primeiro uso explícito de métodos de combinatória poliédrica teria ocorrido em 1943 em um artigo de Rado [127], em que o Lema de Farkas era utilizado de forma indireta [48]. Contudo, o primeiro período de desenvolvimento da combinatória poliédrica teve seu início na década de 50, após o desenvolvimento do método simplex para programação linear. É neste período que surgem as primeiras evidências de que muitos problemas de otimização combinatória (em especial, vários problemas de fluxos em redes) podem ser formulados como problemas de programação linear inteira. Mais ainda, devido a total-unimodularidade de suas matrizes de restrições, tais problemas podiam ser tratados simplesmente como problemas de programação linear. Essas observações constituíram um marco e passaram a direcionar muitas

das pesquisas realizadas nessa época.

Já o segundo período tem seu início em meados da década de 60, quando Edmonds [46] obtém uma descrição linear completa (contudo de tamanho exponencial) para o problema de emparelhamento em um grafo. Uma constatação feita neste período foi a de que mesmo descrições de tamanho exponencial eram úteis no desenvolvimento de algoritmos eficientes para resolução de problemas de otimização combinatória. Este período foi marcado por inúmeros esforços sem sucesso de se obter descrições lineares completas de problemas que mais tarde foram provados ser \mathcal{NP} -difíceis.

Por fim, o terceiro período teve seu início em 1979, com o desenvolvimento do método do elipsóide para programação linear [90]. O resultado fundamental para a combinatória poliédrica não se encontra ligado somente à constatação de que problemas de programação linear podem ser resolvidos em tempo polinomial, mas sim de que é possível resolver eficientemente um problema de programação linear desde que se saiba resolver eficientemente um problema de separação e vice-versa. Essa “equivalência entre otimização e separação” foi uma descoberta que se tornou um marco no tocante ao desenvolvimento de algoritmos na área da combinatória poliédrica. Tornou-se evidente a partir daí que associar poliedros a problemas de otimização combinatória e resolver o problema de separação correspondente poderia ser uma estratégia para se conseguir algoritmos polinomiais para certos problemas [48]. No caso de problemas \mathcal{NP} -difíceis, tal abordagem demonstrou ser interessante, principalmente para problemas de grande porte.

Vale lembrar aqui a evolução algorítmica da abordagem poliédrica na resolução de instâncias reais de problemas \mathcal{NP} -difíceis. O artigo de Dantzig, Fulkerson e Johnson [40] de 1954 é o primeiro a utilizar uma abordagem poliédrica na resolução de um problema \mathcal{NP} -difícil. Tal artigo relata a resolução de uma instância do problema do caixeiro viajante de 49 cidades e ilustra o uso de tal abordagem através de uma série de exemplos, em que muitos dos passos (para se encontrar *planos-de-corte*) são realizados manualmente, de forma interativa e *ad-hoc*. Essa abordagem não foi mais explorada até que em 1977 Grötschel em sua tese de doutoramento [61] realizou um estudo aprofundado da estrutura facial dos politopos associados ao problema do caixeiro viajante simétrico e assimétrico (conseguindo resolver uma instância do problema simétrico com 120 cidades). Contudo, este trabalho implementou uma abordagem poliédrica de forma sistemática e totalmente automática (sem nenhuma intervenção manual para se encontrar *planos-de-corte*).

Na década de 80, vários trabalhos passaram a explorar a abordagem poliédrica na resolução de problemas \mathcal{NP} -difíceis e muitos deles enfocavam, ainda, o problema do caixeiro viajante [64, 116, 117]. Tal problema continua a representar um desafio aos pesquisadores e tem sido fonte de inspiração para o desenvolvimento de várias técnicas que têm se mostrado úteis na resolução de outros problemas de otimização combinatória [48]. Como mencionado, outros problemas \mathcal{NP} -difíceis também foram explorados através dessa abordagem, tais como: problema do subgrafo acíclico [82], problema da ordem linear [62, 63], problema do corte máximo [12], problema da partição em cliques [66, 67] e problema das mochilas múltiplas [47], entre outros. Uma lista mais recente (contendo inclusive artigos da primeira metade da década de 90) pode ser encontrada em [83].

Para todos esses problemas apenas uma descrição parcial de seus politopos é conhecida e em muitos casos o problema de separação só foi resolvido para classes especiais de facetas. Contudo, isso não impediu o desenvolvimento de algoritmos eficientes para resolução de tais problemas e, para casos especiais dos mesmos, até a obtenção de soluções polinomiais.

Nas subseções seguintes são discutidos os detalhes relativos a proposta de implementação de um algoritmo de *branch-and-cut* para resolução do TGP.

4.2.1 Planos-de-Corte

Os métodos poliédricos, quando aplicados a problemas de programação inteira, se caracterizam por iniciar o processo de resolução através de uma aproximação inicial do fecho inteiro do poliedro associado ao conjunto de soluções viáveis. Tal aproximação representa, na realidade, uma relaxação do problema original (como, por exemplo, a relaxação linear) e sua solução não representa necessariamente uma solução viável do problema original (e, na maioria dos casos, ela sequer é inteira).

De modo a prosseguir com o processo de busca, torna-se necessário melhorar a qualidade da aproximação do fecho inteiro do poliedro. Para tanto, inequações válidas serão utilizadas de forma a excluir a atual solução (que não é ótima para o problema original) do conjunto de soluções viáveis. Tais inequações válidas são, geralmente, denominadas *planos-de-corte* ou simplesmente cortes, uma vez que separam a solução atual das demais soluções do problema (isto é, “cortam” o espaço de solução do problema eliminando a solução indesejada). Portanto, para se implementar um método poliédrico para resolução de um problema, torna-se imprescindível a capacidade de se gerar novos *planos-de-corte* caso a solução atual não seja uma solução viável do problema original. Para tanto, pode-se lançar mão de diversas classes de *planos-de-corte* de aplicação geral (*general purpose cutting planes*).

Os primeiros algoritmos de *planos-de-corte* para problemas de programação inteira e inteira mista foram desenvolvidos por Gomory [58, 59, 60], que também demonstrou que tais algoritmos encontravam uma solução ótima após um número finito de iterações. Infelizmente, na prática a utilização dos *planos-de-corte* de Gomory conduz a um tempo de execução muito grande, além de provocar uma instabilidade numérica no sistema de equações (a convergência só é garantida usando-se de precisão infinita) [48]. Em [38], inequações de cobertura mínima (*minimal cover inequalities*) e inequações de $(1, k)$ -configurações (*$(1, k)$ -configuration inequalities*) foram utilizadas para resolução de problemas de otimização linear 0-1. Estas inequações foram derivadas de facetas dos politopos associados aos problemas de mochila definidos por cada restrição da formulação de programação matemática do problema de otimização 0-1. Tal abordagem foi refinada e generalizada em [72, 153]. Já em [5], *planos-de-corte* para problemas de otimização linear 0-1 mista foram gerados através de um método de *lift-and-project*. Os resultados obtidos foram muito promissores e se demonstrou sua superioridade em relação aos cortes de Gomory. Por fim, os cortes de Fenchel para problemas de otimização inteira foram introduzidos em [24].

Entretanto, tais cortes genéricos (*general purpose cutting planes*) se mostram insuficientes para a resolução de problemas de otimização combinatória sendo necessário, para obtenção de um estratégia de sucesso, o desenvolvimento de cortes para

cada problema em particular [83]. Um método de relativo sucesso para obtenção de *planos-de-corte* para um problema específico é o estudo da estrutura facial do fecho inteiro do poliedro associado ao problema. Cada faceta de tal politopo implica em uma inequação válida para a correspondente formulação de programação inteira. Inequações que induzem facetas do politopo não são dominadas por nenhuma outra inequação válida e, portanto, representam os melhores candidatos a *planos-de-corte*. Contudo, a tarefa de se determinar quais inequações induzem facetas, em geral, não é simples (ver [65, 114, 126]). Além disso deve-se ter em mente que, para sua utilização em algoritmos, inequações válidas (definindo ou não facetas) só são úteis caso seus problemas de separação possam ser resolvidos de forma exata ou aproximados através de heurísticas.

Métodos poliédricos (como, por exemplo, o método de *planos-de-corte*) utilizando de cortes desenvolvidos especificamente para um problema normalmente são interrompidos sem encontrar uma solução ótima. Isto se deve a duas razões distintas. Primeiramente, não se conhece para nenhum problema \mathcal{NP} -difícil uma descrição linear completa de seu politopo. Por outro lado, mesmo que se conheça uma grande classe de facetas para um problema, é muito frequente não se dispor de um algoritmo eficiente para resolução do problema de separação para essa classe de facetas. Independentemente disso, grandes instâncias de problemas de otimização combinatória \mathcal{NP} -difíceis podem e têm sido resolvidas através da utilização conjunta de *planos-de-corte* (preferencialmente, facetas) e esquemas de enumeração sofisticados via algoritmo de *branch-and-cut*.

Neste trabalho, as inequações válidas (e, eventualmente, facetas) foram desenvolvidas a partir de um estudo da estrutura facial do fecho inteiro do poliedro associado às soluções viáveis do **TGP**. Mais ainda, o relacionamento entre tal poliedro e outros já investigados na literatura foi explorado nesse estudo de forma semelhante ao que foi realizado em [144, 157] para o problema de projeto de rede com tolerância a falhas. Dessa forma, a identificação de inequações válidas (e, eventualmente, facetas) pôde ser simplificada, além de se facilitar a implementação do algoritmo de *branch-and-cut*, na medida que os respectivos problemas de separação já tinham sido explorados anteriormente. Cabe destacar que essa abordagem poliédrica através da investigação de subestruturas (associadas a poliedros de subproblemas) do poliedro associado ao problema original vem obtendo um relativo sucesso e, mais recentemente, tem sido explorada em diversos trabalhos (ver, por exemplo, [2, 3, 152]).

4.2.2 Método de *Planos-de-Corte*

Uma vez que se obteve um conjunto de inequações válidas (e, eventualmente, facetas) para o **TGP** conforme descrito acima, passou-se a utilizá-lo na implementação de um método de *planos-de-corte*. Em vista do que foi exposto anteriormente, não se deve esperar que tal procedimento seja capaz de encontrar uma solução ótima para o **TGP**.

De fato, os resultados de experimentos preliminares com cortes genéricos não apresentaram um desempenho satisfatório. Contudo, um método de *planos-de-corte* pode ser visto como o passo inicial de um algoritmo de *branch-and-cut*. Sendo assim, a qualidade da descrição linear obtida pelo estudo da estrutura facial do fecho inteiro

do poliedro de soluções (juntamente com os algoritmos de separação exatos ou não) poderia ser avaliada antes mesmo da implementação completa do procedimento de busca.

Portanto, o desenvolvimento de um método de *planos-de-corte* para o **TGP** foi utilizado para se validar e, eventualmente, rever o conjunto de inequações válidas obtidas para o problema.

4.2.3 Algoritmo de *Cut-and-Branch*

Os algoritmos de *branch-and-cut* podem ser vistos como esquemas de enumeração sofisticados em que estratégias meticolosas de “fixação” de variáveis e ramificação (*branching*) são combinadas para obtenção de um processo de busca eficiente. Além disso, durante o processo de busca devem ser obtidos e utilizados “novos” cortes para cada subproblema a ser resolvido.

Tal processo se caracteriza por necessitar de estruturas “especiais” para o armazenamento e processamento do conjunto de inequações ativas em um dado momento da busca. Sendo assim, talvez seja conveniente se avaliar o real ganho de tal procedimento através da implementação de uma busca “simplificada”.

Nesta abordagem (denominada de algoritmo de *cut-and-branch*), após a aplicação de um método de *planos-de-corte*, um algoritmo de *branch-and-bound* padrão (sem a inclusão de novos cortes) foi utilizado para se obter uma solução ótima do problema.

4.2.4 Algoritmo de Busca

Finalmente, o objetivo principal desta segunda abordagem foi o desenvolvimento e teste de um algoritmo de *branch-and-cut* para o **TGP**.

Conforme foi mencionado, a implementação de tal algoritmo envolve vários detalhes (sendo que muitos deles fogem ao escopo deste trabalho). Sendo assim, para viabilizar a sua realização foi utilizado um *framework* padrão como, por exemplo, o COIN-CBC [78], o ABACUS [150] ou o SYMPHONY [129]. O *framework* utilizado neste trabalho foi o COIN-CBC, uma vez que, além de ser multiplataforma e suportar diversos *LP-solvers*, ele faz parte do esforço da INFORMS em fornecer software (incluindo os fontes) gratuito e de boa qualidade para uso no desenvolvimento e pesquisa em otimização.

4.3 Considerações Finais

Neste capítulo, procurou-se apresentar os principais aspectos das abordagens selecionadas para resolução do **TGP**. Além disso, uma descrição geral de tais abordagens foi realizada com intuito de esclarecer os principais pontos enfocados neste trabalho. Entretanto, diversos detalhes pertinentes as duas abordagens foram omitidos para facilitar a leitura e o entendimento deste documento.

Capítulo 5

Abordagem Lagrangeana do TGP

NESTE capítulo é apresentada uma abordagem para resolução do **TGP** através da relaxação lagrangeana. Na próxima seção são apresentados limites inferiores obtidos a partir da relaxação de diferentes conjuntos de restrições da formulação do problema TGP_C apresentada na subseção 3.1.4. Na seção 5.2 a heurística utilizada para geração de limites superiores para o **TGP** é apresentada. Em seguida, as seções 5.3, 5.4 e 5.5 discutem aspectos relevantes da implementação do método de subgradiente (utilizado na resolução do problema lagrangeano dual), do algoritmo de *branch-and-bound* e da heurística lagrangeana, respectivamente. Já a seção 5.6 apresenta os resultados numéricos obtidos (e apresentados parcialmente em [119, 120]); enquanto que uma comparação com outras abordagens encontradas na literatura é apresentada na seção 5.7, seguida por algumas considerações finais na seção 5.8.

5.1 Limites Inferiores para TGP

Nesta seção a relaxação lagrangeana é utilizada para obtenção de quatro limites inferiores para a formulação do problema TGP_C apresentada na subseção 3.1.4. Tais limites são utilizados mais adiante (ver seções 5.3 e 5.4) na implementação de um procedimento de busca – *branch-and-bound* – para a resolução do problema.

De modo a se realizar a relaxação lagrangeana da formulação proposta, um vetor de multiplicadores de Lagrange π , tal que $\pi_i^p \in \mathbb{R}, \forall t_p \in P, \forall i \in N_\ell$, é associado às restrições de conservação de fluxo dos nós – restrições (3.10b). Já as restrições (3.10c) são associadas a um vetor de multiplicadores ϕ , tal que $\phi_i \in \mathbb{R}, \forall i \in N_\ell^o$. E, finalmente, as restrições que limitam a capacidade dos arcos – restrições (3.10d), são associadas a um vetor de multiplicadores de Lagrange λ , tal que $\lambda_\ell \in \mathbb{R}_+, \forall \ell \in A_\ell$.

Posto isto, são apresentados a seguir as formulações e procedimentos de cálculo de quatro limites inferiores obtidos através do uso da relaxação lagrangeana. No primeiro deles, todas as restrições mencionadas acima são relaxadas; enquanto que no segundo limite inferior, são relaxadas apenas as restrições (3.10d) responsáveis por estabelecer uma ligação entre as variáveis de fluxo (f_ℓ^p) e as variáveis de decisão (w_ℓ) e estabelecer capacidades para os arcos. Já no terceiro limite inferior são relaxadas as restrições (3.10d) juntamente com as restrições de conservação de fluxo

(3.10b). E, finalmente, no quarto limite inferior as restrições (3.10d) são relaxadas juntamente com as restrições de conservação de *containers virtuais* (3.10c).

5.1.1 Relaxação #1

No primeiro limite inferior para a formulação TGP_C , as restrições de (3.10b) a (3.10d) são relaxadas, permanecendo apenas as restrições que definem as variáveis de fluxo e de decisão – restrições (3.10e), (3.10f) e (3.10g). Dessa forma, o primeiro limite inferior, L_1^{TGP} , pode ser formulado da seguinte maneira:

$$L_1^{\text{TGP}}(\pi, \phi, \lambda) = \min \sum_{\ell \in A_\ell} c_\ell w_\ell + \sum_{t_p \in P} \sum_{i \in N_\ell} \pi_i^p \left(\sum_{\ell \in A_\ell^+(i)} f_\ell^p - \sum_{\ell \in A_\ell^-(i)} f_\ell^p - b_i^p \right) + \sum_{i \in N_\ell^o} \phi_i \left(\sum_{\ell \in A_\ell^+(i)} w_\ell - \sum_{\ell \in A_\ell^-(i)} w_\ell \right) + \sum_{\ell \in A_\ell} \lambda_\ell \left(\sum_{t_p \in P} \delta_p f_\ell^p - \Delta w_\ell \right)$$

sujeito a: (3.10e), (3.10f) e (3.10g).

Rearranjando os termos acima, pode-se reescrever o primeiro limite inferior como se segue:

$$L_1^{\text{TGP}}(\pi, \phi, \lambda) = \min \sum_{\ell \in A_\ell} \tilde{c}_\ell w_\ell + \sum_{t_p \in P} \sum_{\ell \in A_\ell} \tilde{c}_\ell^p f_\ell^p + \sum_{t_p \in P} (\pi_{d_p}^p - \pi_{s_p}^p) f_p \quad (5.1)$$

sujeito a: (3.10e), (3.10f) e (3.10g).

em que os valores dos custos modificados \tilde{c}_ℓ e \tilde{c}_ℓ^p são dados por:

$$\tilde{c}_\ell = \begin{cases} c_\ell - \Delta \lambda_\ell + \phi_i & , \ell = (i, j) \mid i \in N_\ell^o, j \in N_\ell^e \\ c_\ell - \Delta \lambda_\ell - \phi_j & , \ell = (i, j) \mid i \in N_\ell^e, j \in N_\ell^o \\ c_\ell - \Delta \lambda_\ell + \phi_i - \phi_j & , \ell = (i, j) \mid i \in N_\ell^o, j \in N_\ell^o \end{cases} \quad (5.2)$$

$$\tilde{c}_\ell^p = \delta_p \lambda_\ell + \pi_i^p - \pi_j^p \quad , \ell = (i, j) \quad (5.3)$$

De modo a se viabilizar o cálculo desse limite inferior, além de se obter uma solução relaxada mais próxima da efetiva solução do problema, pode-se agregar as restrições (5.4) à formulação dada anteriormente de modo que o valor do fluxo de um dado produto através de um arco seja limitado pelo número máximo de unidades desse produto que podem trafegar pelo mesmo.

$$f_\ell^p \leq m_p \quad , \forall t_p \in P, \forall \ell \in A_\ell \quad (5.4)$$

Em verdade, pode-se considerar que tais restrições já pertenciam à formulação original do problema – sem contudo modificar sua solução – e que foram mantidas intocadas no modelo relaxado. De todo modo, a resolução dessa formulação “aumentada” pode ser obtida através do seguinte procedimento em que \tilde{f}_ℓ^p e \tilde{w}_ℓ representam os valores das variáveis de fluxo e de decisão, respectivamente:

Procedimento de Cálculo do Limite Inferior L_1^{TGP}

Passo 1 para $\forall t_p \in P$ faça
 para $\forall \ell \in A_\ell$ faça
 se $\tilde{c}_\ell^p < 0$ então
 faça $\bar{f}_\ell^p = m_p$
 senão
 faça $\bar{f}_\ell^p = 0$
 fimse
 fimpara
fimpara

Passo 2 para $\forall \ell \in A_\ell$ faça
 calcule $\hat{c}_\ell = \tilde{c}_\ell + \sum_{t_p \in P | \tilde{c}_\ell^p < 0} \tilde{c}_\ell^p m_p$
 se $\hat{c}_\ell < 0$ então
 faça $\bar{w}_\ell = 1$
 senão
 faça $\bar{w}_\ell = 0$
 fimse
fimpara

Desse modo o valor do primeiro limite inferior, L_1^{TGP} , é dado por:

$$L_1^{\text{TGP}}(\pi, \phi, \lambda) = \sum_{\ell \in A_\ell} \hat{c}_\ell \bar{w}_\ell + \sum_{t_p \in P} (\pi_{d_p}^p - \pi_{s_p}^p) f_p. \quad (5.5)$$

5.1.2 Relaxação #2

Já no segundo limite inferior, L_2^{TGP} , são relaxadas apenas as restrições (3.10d). Isto resulta na eliminação do relacionamento entre as variáveis de fluxo e de decisão. Mais ainda, as variáveis de fluxo de produtos distintos também deixam de ser interdependentes uma vez que essas restrições – restrições (3.10d) – eram responsáveis pelo limite sobre o volume total de produtos transportados através de um arco.

Dessa forma, o segundo limite inferior, L_2^{TGP} , pode ser formulado da seguinte maneira:

$$L_2^{\text{TGP}}(\lambda) = \min \sum_{\ell \in A_\ell} c_\ell w_\ell + \sum_{\ell \in A_\ell} \lambda_\ell \left(\sum_{t_p \in P} \delta_p f_\ell^p - \Delta w_\ell \right)$$

sujeito a: (3.10b), (3.10c), (3.10e), (3.10f) e (3.10g).

Rearranjando os termos acima, pode-se reescrever o segundo limite inferior como dois subproblemas disjuntos da seguinte forma:

$$L_2^{\text{TGP}}(\lambda) = L_{2W}^{\text{TGP}}(\lambda) + L_{2F}^{\text{TGP}}(\lambda) \quad (5.6)$$

em que:

$$L_{2W}^{\text{TGP}}(\lambda) = \min \sum_{\ell \in A_\ell} (c_\ell - \Delta \lambda_\ell) w_\ell \quad L_{2F}^{\text{TGP}}(\lambda) = \min \sum_{t_p \in P} \sum_{\ell \in A_\ell} \delta_p \lambda_\ell f_\ell^p$$

sujeito a: (3.10c) e (3.10f).

sujeito a: (3.10b), (3.10e) e (3.10g).

O primeiro subproblema, L_{2W}^{TGP} , deve ser solucionado através da resolução de um problema de fluxo de custo mínimo. Antes porém, a semelhança do que foi feito no cálculo do primeiro limite inferior (ver subseção 5.1.1) para se obter uma solução relaxada mais próxima da efetiva solução do problema, deve-se agregar as restrições (5.7a) e (5.7b) – inequações válidas – à formulação do subproblema L_{2W}^{TGP} dada anteriormente.

$$\sum_{\ell \in A_{\ell}^{+}(i)} w_{\ell} \geq \beta_i, \forall i \in N_{\ell}^e \quad (5.7a)$$

$$\sum_{\ell \in A_{\ell}^{-}(i)} w_{\ell} \geq \gamma_i, \forall i \in N_{\ell}^e \quad (5.7b)$$

em que β_i e $\gamma_i, \forall i \in N_{\ell}^e$, representam o número mínimo de arcos necessários para se transportar, respectivamente, todo o tráfego originado e destinado a cada nó de *add-drop* e são dados pelas seguintes expressões:

$$\beta_i = \left[\sum_{t_p \in P \mid s_p = i} \frac{f_p}{m_p} \right], \forall i \in N_{\ell}^e, \quad \gamma_i = \left[\sum_{t_p \in P \mid d_p = i} \frac{f_p}{m_p} \right], \forall i \in N_{\ell}^e.$$

Desse modo assegura-se que o número de arcos utilizados na solução obtida seja suficiente para transportar toda a oferta (ou demanda) de um produto originado (ou destinado) a um nó de *add-drop*. Em verdade, como antes, pode-se considerar que tais restrições já pertenciam à formulação original do problema – sem contudo modificar sua solução – e que foram mantidas intocadas no modelo relaxado.

Como mencionado anteriormente, o subproblema $L_{2W}^{\text{TGP}}(\lambda)$ pode ser solucionado através da resolução de um problema de fluxo de custo mínimo. Nessa abordagem as facilidades instaladas nos arcos, isto é, os *containers virtuais* serão tratados como um único tipo de produto que deve fluir entre os nós de oferta e demanda através de caminhos disjuntos, de modo a prevenir o *waveclash* (utilização simultânea da mesma facilidade no transporte de diferentes conjuntos de produtos através de um arco). Para tanto, a representação em camadas modificada descrita ao final da subseção 3.1.3 deve ser utilizada, de modo que se possa representar simultaneamente tanto a oferta quanto a demanda de facilidades (*containers virtuais*) associadas a um nó de *add-drop*, conforme descrito pelas restrições (5.7a) e (5.7b). Sendo assim, a oferta de facilidades descrita por (5.7a) será associada aos nós de *add*, enquanto que aos nós de *drop* serão associadas as demandas descritas por (5.7b).

Seja $\tilde{G}_{\ell} = (\tilde{N}_{\ell}, \tilde{A}_{\ell})$ a representação em camadas modificada construída de acordo com o procedimento descrito ao final da subseção 3.1.3. Em tal representação o conjunto de nós, \tilde{N}_{ℓ} , pode ser particionado em três subconjuntos disjuntos, isto é, $\tilde{N}_{\ell} = \tilde{N}_{\ell}^a \cup \tilde{N}_{\ell}^d \cup \tilde{N}_{\ell}^o$, em que \tilde{N}_{ℓ}^a , \tilde{N}_{ℓ}^d e \tilde{N}_{ℓ}^o representam, respectivamente, os conjuntos de nós de *add*, nós de *drop* e nós óticos.

Além disso, de modo a tratar uma eventual diferença existente entre a oferta total de facilidades, $\beta_T = \sum_{i \in \tilde{N}_{\ell}^a} \beta_i$, e a demanda total, $\gamma_T = \sum_{i \in \tilde{N}_{\ell}^d} \gamma_i$, deve-se acrescentar um nó artificial, η , à representação em camadas modificada. Para todo nó de *add* $i \in \tilde{N}_{\ell}^a$, deverá ser acrescentado um arco partindo do nó artificial η em direção a i . Analogamente, para todo nó de *drop* $i \in \tilde{N}_{\ell}^d$, deverá ser acrescentado um

arco partindo do nó de *drop* i em direção ao nó artificial η . A oferta (ou demanda) associada ao nó artificial η será dada por $\gamma_T - \beta_T$, fazendo com que tal nó artificial atue como nó de *add* caso haja um excesso de demanda, ou como nó de *drop* quando a oferta total exceder a demanda total.

Dessa forma, seja $\tilde{A}_\ell^+(i)$ o conjunto de todos arcos que saem de um nó $i \in (\tilde{N}_\ell \cup \{\eta\})$, enquanto que $\tilde{A}_\ell^-(i)$ representa o conjunto de todos os arcos que chegam em um nó $i \in (\tilde{N}_\ell \cup \{\eta\})$. Portanto:

$$\tilde{A}_\ell^+(i) = \begin{cases} \{\ell = (i, j) \mid j \in \tilde{N}_\ell\} & , \forall i \in (\tilde{N}_\ell^a \cup \tilde{N}_\ell^o) \\ \{\ell = (i, j) \mid j \in (\tilde{N}_\ell \cup \{\eta\})\} & , \forall i \in \tilde{N}_\ell^d \\ \{\ell = (\eta, j) \mid j \in \tilde{N}_\ell^a\} & , i = \eta \end{cases} \quad (5.8)$$

$$\tilde{A}_\ell^-(i) = \begin{cases} \{\ell = (j, i) \mid j \in \tilde{N}_\ell\} & , \forall i \in (\tilde{N}_\ell^d \cup \tilde{N}_\ell^o) \\ \{\ell = (j, i) \mid j \in (\tilde{N}_\ell \cup \{\eta\})\} & , \forall i \in \tilde{N}_\ell^a \\ \{\ell = (j, \eta) \mid j \in \tilde{N}_\ell^d\} & , i = \eta \end{cases} \quad (5.9)$$

Dado o conjunto de arcos artificiais $\tilde{A}_\ell^\eta = \{(\eta, j) \mid j \in \tilde{N}_\ell^a\} \cup \{(i, \eta) \mid i \in \tilde{N}_\ell^d\}$, para os quais os custos são nulos, deve-se, ainda, estender o conjunto de variáveis, de modo a se associar a cada arco artificial $\ell \in \tilde{A}_\ell^\eta$ uma variável w_ℓ inteira não negativa. Posto isto, o subproblema $L_{2W}^{\text{TGP}}(\lambda)$ pode ser reformulado da seguinte forma:

$$L_{2W}^{\text{TGP}}(\lambda) = \min \sum_{\ell \in A_\ell} (c_\ell - \Delta\lambda_\ell) w_\ell \quad (5.10a)$$

sujeito a:

$$\sum_{\ell \in \tilde{A}_\ell^+(i)} w_\ell - \sum_{\ell \in \tilde{A}_\ell^-(i)} w_\ell = b_i^w \quad , \forall i \in (\tilde{N}_\ell \cup \{\eta\}) \quad (5.10b)$$

$$w_\ell \leq 1 \quad , \forall \ell \in \tilde{A}_\ell \quad (5.10c)$$

$$w_\ell \geq 0 \quad , \forall \ell \in (\tilde{A}_\ell \cup \tilde{A}_\ell^\eta) \quad (5.10d)$$

$$w_\ell \text{ inteiro} \quad , \forall \ell \in (\tilde{A}_\ell \cup \tilde{A}_\ell^\eta) \quad (5.10e)$$

em que b_i^w é dado por:

$$b_i^w = \begin{cases} \beta_i & , \text{ se } i \in \tilde{N}_\ell^a \\ -\gamma_i & , \text{ se } i \in \tilde{N}_\ell^d \\ 0 & , \text{ se } i \in \tilde{N}_\ell^o \\ \gamma_T - \beta_T & , \text{ se } i = \eta \end{cases} \quad , \forall i \in (\tilde{N}_\ell \cup \{\eta\}). \quad (5.11)$$

As restrições de integralidade (5.10e) podem ser desprezadas, uma vez que a matriz de coeficientes do conjunto restante de restrições, (5.10b)–(5.10d), é totalmente unimodular. Dessa forma, a solução do primeiro subproblema, L_{2W}^{TGP} , pode ser obtida através da resolução do problema de fluxo de custo mínimo representado (5.10a)–(5.10d).

Considere que $\bar{w}_\ell, \forall \ell \in (\tilde{A}_\ell \cup \tilde{A}_\ell^\eta)$, representam os valores das variáveis de decisão obtidos a partir da resolução do problema de fluxo de custo mínimo (5.10a)–(5.10d). Desse modo o valor da solução do primeiro subproblema, L_{2W}^{TGP} , é dado por:

$$L_{2W}^{\text{TGP}}(\lambda) = \sum_{\ell \in A_\ell} (c_\ell - \Delta\lambda_\ell) \bar{w}_\ell. \quad (5.12)$$

Por sua vez, o segundo subproblema, $L_{2F}^{\text{TGP}}(\lambda)$, definido anteriormente pode ser decomposto em $n_p = |P|$ problemas independentes de modo que:

$$L_{2F}^{\text{TGP}}(\lambda) = \sum_{t_p \in P} \delta_p L_{2F_p}^{\text{TGP}}(\lambda) \quad (5.13)$$

em que para cada produto $t_p \in P$:

$$L_{2F_p}^{\text{TGP}}(\lambda) = \min \sum_{\ell \in A_\ell} \lambda_\ell f_\ell^p \quad (5.14a)$$

sujeito a:

$$\sum_{\ell \in A_\ell^+(i)} f_\ell^p - \sum_{\ell \in A_\ell^-(i)} f_\ell^p = b_i^p, \quad \forall i \in N_\ell \quad (5.14b)$$

$$f_\ell^p \geq 0, \quad \forall \ell \in A_\ell \quad (5.14c)$$

$$f_\ell^p \text{ inteiro}, \quad \forall \ell \in A_\ell \quad (5.14d)$$

De maneira análoga ao subproblema $L_{2W}^{\text{TGP}}(\lambda)$, as restrições (5.14d) – restrições de integralidade – podem ser desprezadas, uma vez que a matriz de coeficientes do conjunto restante de restrições, (5.14b)–(5.14c), é totalmente unimodular. Além disso, cada um dos subproblemas $L_{2F_p}^{\text{TGP}}(\lambda)$ pode ser resolvido através do cálculo do caminho mínimo entre o nó de origem s_p e o nó de destino d_p de cada produto.

Define-se \mathcal{C}^p como o conjunto de arcos pertencentes ao caminho mínimo entre s_p e d_p calculado utilizando-se λ_ℓ como custo de cada arco. Sendo assim, a solução do segundo subproblema pode ser obtida roteando-se o tráfego de cada produto $t_p \in P$ através do caminho mínimo \mathcal{C}^p .

Considere que $\bar{f}_\ell^p, \forall \ell \in A_\ell, \forall t_p \in P$, representam os valores obtidos, dessa maneira, para as variáveis de fluxo. Desse modo o valor da solução do segundo subproblema, L_{2F}^{TGP} , é dado por:

$$L_{2F}^{\text{TGP}}(\lambda) = \sum_{t_p \in P} \delta_p \sum_{\ell \in A_\ell} \lambda_\ell \bar{f}_\ell^p = \sum_{t_p \in P} \delta_p \sum_{\ell \in \mathcal{C}^p} (\lambda_\ell f_p) = \sum_{t_p \in P} \delta_p f_p \left(\sum_{\ell \in \mathcal{C}^p} \lambda_\ell \right). \quad (5.15)$$

Portanto, o valor total do segundo limite inferior é dado por:

$$L_2^{\text{TGP}}(\lambda) = L_{2W}^{\text{TGP}}(\lambda) + L_{2F}^{\text{TGP}}(\lambda) = \sum_{\ell \in A_\ell} (c_\ell - \Delta \lambda_\ell) \bar{w}_\ell + \sum_{t_p \in P} \delta_p f_p \left(\sum_{\ell \in \mathcal{C}^p} \lambda_\ell \right). \quad (5.16)$$

Cumpramos ressaltar que, apesar do cálculo de L_2^{TGP} envolver a resolução de um problema de fluxo de custo mínimo, uma implementação “cuidadosa” do mesmo pode tirar proveito do fato de que apenas os custos dos arcos são modificados entre duas iterações distintas. Sendo assim, uma vez calculada uma primeira solução do subproblema de fluxo de custo mínimo, as demais soluções podem ser obtidas através da reotimização da solução anterior que continua sendo viável e, eventualmente, até mesmo ótima, apesar de uma eventual alteração no valor dos custos.

5.1.3 Relaxação #3

Já no terceiro limite inferior, L_3^{TGP} , são relaxadas as restrições (3.10d), juntamente com as restrições de conservação de fluxo (3.10b).

Dessa forma, o terceiro limite inferior, L_3^{TGP} , pode ser formulado da seguinte maneira:

$$L_3^{\text{TGP}}(\pi, \lambda) = \min \sum_{\ell \in A_\ell} c_\ell w_\ell + \sum_{t_p \in P} \sum_{i \in N_\ell} \pi_i^p \left(\sum_{\ell \in A_\ell^+(i)} f_\ell^p - \sum_{\ell \in A_\ell^-(i)} f_\ell^p - b_i^p \right) + \sum_{\ell \in A_\ell} \lambda_\ell \left(\sum_{t_p \in P} \delta_p f_\ell^p - \Delta w_\ell \right)$$

sujeito a: (3.10c), (3.10e), (3.10f) e (3.10g).

Rearranjando os termos acima, pode-se reescrever o terceiro limite inferior como dois subproblemas disjuntos da seguinte forma:

$$L_3^{\text{TGP}}(\pi, \lambda) = L_{3W}^{\text{TGP}}(\lambda) + L_{3F}^{\text{TGP}}(\pi, \lambda) + \sum_{t_p \in P} (\pi_{d_p}^p - \pi_{s_p}^p) f_p \quad (5.17)$$

em que:

$$L_{3W}^{\text{TGP}}(\lambda) = \min \sum_{\ell \in A_\ell} (c_\ell - \Delta \lambda_\ell) w_\ell \quad L_{3F}^{\text{TGP}}(\pi, \lambda) = \sum_{t_p \in P} \sum_{\ell \in A_\ell} \tilde{c}_\ell^p f_\ell^p$$

sujeito a: (3.10c) e (3.10f).

sujeito a: (3.10e) e (3.10g).

e \tilde{c}_ℓ^p é dado pela equação (5.3).

O primeiro subproblema, L_{3W}^{TGP} , é equivalente ao primeiro subproblema da relaxação anterior (L_{2W}^{TGP}) e deve ser solucionado da mesma forma, isto é, através do acréscimo das inequações válidas (5.7a) e (5.7b) e resolução do problema de fluxo de custo mínimo associado (ver subseção 5.1.2).

Já o segundo subproblema, L_{3F}^{TGP} , pode ser resolvido por inspeção a semelhança do que foi descrito para cálculo do primeiro limite inferior (ver subseção 5.1.1). De fato, como no primeiro limite inferior, pode-se adotar as restrições (5.4) e utilizar o **Passo 1** do algoritmo descrito para o cálculo do primeiro limite inferior para se solucionar o subproblema L_{3F}^{TGP} .

Seja $(\bar{f}_\ell^p, \bar{w}_\ell)$ a solução obtida da forma descrita acima e A_ℓ^p o conjunto de arcos $\ell \in A_\ell$ para os quais o custo modificado \tilde{c}_ℓ^p associado ao produto $t_p \in P$ é negativo, isto é, $A_\ell^p = \{\ell \in A_\ell \mid \tilde{c}_\ell^p < 0\}, \forall t_p \in P$. Sendo assim, o valor do terceiro limite inferior, L_3^{TGP} , é dado por:

$$L_3^{\text{TGP}}(\pi, \lambda) = \sum_{\ell \in A_\ell} (c_\ell - \Delta \lambda_\ell) \bar{w}_\ell + \sum_{t_p \in P} m_p \left(\sum_{\ell \in A_\ell^p} \tilde{c}_\ell^p \right) + \sum_{t_p \in P} (\pi_{d_p}^p - \pi_{s_p}^p) f_p. \quad (5.18)$$

De forma análoga ao descrito para o segundo limite inferior, na resolução do subproblema de fluxo de custo mínimo, L_{3W}^{TGP} , pode-se utilizar de reotimização de modo a se reaproveitar a solução anterior na obtenção de uma nova.

5.1.4 Relaxação #4

Finalmente, no quarto limite inferior, L_4^{TGP} , são relaxadas as restrições (3.10d), juntamente com as restrições de conservação de *containers virtuais* (3.10c).

Sendo assim, o quarto limite inferior, L_4^{TGP} , pode ser formulado da seguinte maneira:

$$L_4^{\text{TGP}}(\phi, \lambda) = \min \sum_{\ell \in A_\ell} c_\ell w_\ell + \sum_{i \in N_\ell^o} \phi_i \left(\sum_{\ell \in A_\ell^+(i)} w_\ell - \sum_{\ell \in A_\ell^-(i)} w_\ell \right) + \sum_{\ell \in A_\ell} \lambda_\ell \left(\sum_{t_p \in P} \delta_p f_\ell^p - \Delta w_\ell \right)$$

sujeito a: (3.10b), (3.10e), (3.10f) e (3.10g).

Rearranjando os termos acima, pode-se reescrever o quarto limite inferior como dois subproblemas disjuntos da seguinte forma:

$$L_4^{\text{TGP}}(\phi, \lambda) = L_{4W}^{\text{TGP}}(\phi, \lambda) + L_{4F}^{\text{TGP}}(\lambda) \quad (5.19)$$

em que:

$$L_{4W}^{\text{TGP}}(\phi, \lambda) = \min \sum_{\ell \in A_\ell} \tilde{c}_\ell w_\ell \quad L_{4F}^{\text{TGP}}(\lambda) = \min \sum_{t_p \in P} \sum_{\ell \in A_\ell} \delta_p \lambda_\ell f_\ell^p$$

sujeito a: (3.10f).

sujeito a: (3.10b), (3.10e) e (3.10g).

e \tilde{c}_ℓ é dado pela equação (5.2).

O segundo subproblema, L_{4F}^{TGP} , é equivalente ao segundo subproblema da segunda relaxação (L_{2F}^{TGP}) e deve ser solucionado da mesma forma, isto é, através do cálculo de caminhos mínimos (ver subsecção 5.1.2).

Já o primeiro subproblema, L_{4W}^{TGP} , pode ser resolvido por inspeção, bastando fazer w_ℓ igual a um para todo arco $\ell \in A_\ell$ cujo custo modificado \tilde{c}_ℓ for negativo.

Seja (f_ℓ^p, \bar{w}_ℓ) a solução obtida da forma descrita acima e $A_\ell^<$ o conjunto de arcos $\ell \in A_\ell$ para os quais o custo modificado \tilde{c}_ℓ é negativo, isto é, $A_\ell^< = \{\ell \in A_\ell \mid \tilde{c}_\ell < 0\}$. Sendo assim, o valor do quarto limite inferior, L_4^{TGP} , é dado por:

$$L_4^{\text{TGP}}(\phi, \lambda) = \sum_{\ell \in A_\ell^<} \tilde{c}_\ell + \sum_{t_p \in P} \delta_p f_p \left(\sum_{\ell \in C^p} \lambda_\ell \right). \quad (5.20)$$

5.2 Limite Superior para TGP

De modo a se implementar um procedimento de busca – *branch-and-bound* – para a resolução do problema, faz-se necessária a obtenção de uma solução viável (limite superior) para o mesmo. Mais ainda, tal limite superior deve ser atualizado, levando-se em conta, se possível, os limites inferiores calculados até então. Tal procedimento visa obter soluções melhores (mais baratas) para o problema, além de possibilitar uma redução do espaço de busca e, conseqüentemente, uma melhora do desempenho do algoritmo.

Para tanto, desenvolveu-se uma heurística que utiliza o conjunto de facilidades (arcos da representação em camadas) selecionadas durante o cálculo do limite inferior como base para obtenção de uma solução viável. Em verdade, tal heurística representa uma adaptação do método proposto por Busacker-Gowen para a resolução de problemas de fluxo de custo mínimo de um único produto através do cálculo de sucessivos caminhos mínimos [25, 79, 81]. Heurísticas similares têm sido propostas por outros autores para o problema de fluxo de custo mínimo de vários produtos [13, 73] e, até mesmo, para o problema de projeto/planejamento de redes capacitadas [74]. Deve-se destacar que, tanto para o segundo quanto para o terceiro limite inferior, o conjunto de facilidades selecionadas atende parcialmente o conjunto original de restrições, tornando mais simples, dessa forma, se obter boas soluções viáveis. Contudo, para o primeiro limite inferior (bem como para o último), o mesmo não se verifica.

Seja \hat{A}_0 o conjunto de facilidades (arcos) disponíveis para o cálculo do limite superior. Inicialmente esse conjunto pode ser igual ao conjunto de arcos \bar{A}_ℓ selecionados durante o cálculo do limite inferior, isto é, $\hat{A}_0 = \{\ell \in A_\ell \mid \bar{w}_\ell = 1\}$, em que \bar{w}_ℓ representa o valor da variável de decisão associada ao arco $\ell \in A_\ell$ obtido pelo cálculo do limite inferior, ou, até mesmo, através de uma outra heurística qualquer como, por exemplo, $\hat{A}_0 = \{\ell \in A_\ell \mid \bar{f}_\ell^p > 0\}$ ou $\hat{A}_0 = \{\ell \in A_\ell \mid \bar{w}_\ell = 1 \text{ ou } \bar{f}_\ell^p > 0\}$.

A primeira parte da heurística roteia inicialmente cada produto através do caminho mais curto entre sua origem e seu destino, utilizando-se apenas dos arcos pertencentes a \bar{A}_ℓ . Tal roteamento é realizado em ordem decrescente de ocupação da capacidade de transporte, isto é, o produto cuja razão f^p/m_p for maior é roteado primeiro, sendo seguido por aquele que apresentar o segundo maior valor e assim sucessivamente. Durante tal procedimento a capacidade de cada arco do caminho utilizado no roteamento – inicialmente igual a Δ – é atualizada, isto é, reduzida de modo a refletir a diminuição da capacidade de transporte em decorrência do roteamento (mesmo que parcial) do tráfego.

De fato, para cada produto é calculado o caminho de custo mínimo entre sua fonte e seu sumidouro que ainda permita o transporte de algum tráfego, isto é, a capacidade disponível do arco mais saturado (ocupado) do caminho ainda é maior que zero. Feito isto, o produto é, então, roteado através desse caminho. Caso a capacidade de transporte disponível no caminho não seja suficiente para se rotear todo o tráfego do produto, um novo caminho é calculado e utilizado. Essa sequência de operações é repetida até que todo o tráfego de um produto tenha sido roteado ou que não se encontre um caminho de capacidade disponível entre a fonte e o sumidouro de um produto. Nesse último caso, considera-se que o conjunto de facilidades (arcos) selecionadas até o momento não é suficiente para permitir a obtenção de uma solução viável e a heurística é interrompida sem obter uma solução viável.

Caso a primeira parte da heurística tenha sucesso em obter um roteamento inicial para cada produto, a solução encontrada não é necessariamente viável, uma vez que a quantidade de arcos alocados que saem de um nó ótico pode ser diferente da quantidade de arcos alocados que chegam no mesmo nó, violando, assim, as restrições (3.10c). Sendo assim, de modo a tornar viável a solução obtida na primeira parte da heurística, verifica-se se cada nó ótico atende ou não aos limites impostos pelas restrições (3.10c).

Para tanto, basta observar que o número de facilidades que saem de um nó ótico de entrada deve ser igual a um, uma vez que, por construção, seu grau de entrada é um. Caso se encontre um nó ótico de entrada que viole essa restrição, basta fazer com que todo seu tráfego seja roteado através do correspondente nó de *add-drop*, antes de seguir para seu destino. Uma linha de raciocínio análoga pode ser seguida para os nós óticos de saída, podendo-se adotar uma solução semelhante.

Como exposto, essa heurística consiste na resolução de vários problemas de caminho mínimo e, portanto, é muito rápida. Contudo, apesar de produzir soluções viáveis, os resultados obtidos de sua aplicação podem não ser satisfatórios (dado a excessiva utilização de nós de *add-drop* na viabilização do roteamento inicial, por exemplo), sendo portanto necessária a adoção de heurísticas complementares. Uma vez obtida uma solução viável, tais heurísticas complementares buscariam melhorar a qualidade do limite superior gerado (isto é, reduzir seu custo) através, por exemplo, da “combinação” de *lightpaths*, mantendo a viabilidade da solução encontrada durante todo o processo.

Seja $(\hat{w}_\ell, \hat{f}_\ell^p)$ a solução viável obtida ao final de todo esse processo e \hat{A}_ℓ o conjunto de arcos selecionados, isto é, $\hat{A}_\ell = \{\ell \in A_\ell \mid \hat{w}_\ell = 1\}$. Dessa forma, o valor do limite superior, \hat{L} , é dado por:

$$\hat{L} = \sum_{\ell \in \hat{A}_\ell} c_\ell. \quad (5.21)$$

5.3 Método de Subgradiente para TGP

Para cada conjunto de multiplicadores de Lagrange, obtém-se um limite inferior para o TGP, dado por $L_1^{\text{TGP}}(\pi, \phi, \lambda)$, $L_2^{\text{TGP}}(\lambda)$, $L_3^{\text{TGP}}(\pi, \lambda)$ ou $L_4^{\text{TGP}}(\phi, \lambda)$. De modo a se obter uma solução para o problema, deve-se procurar maximizar o valor do limite inferior obtido.

Neste sentido, o melhor limite inferior é dado pela solução do problema lagrangeano dual P_1^D , caso a relaxação #1 seja aplicada:

$$(P_1^D) \quad \max_{\substack{\pi, \phi \in \mathbb{R} \\ \lambda \geq 0}} L_1^{\text{TGP}}(\pi, \phi, \lambda), \quad (5.22)$$

caso se utilize a relaxação #2, o problema lagrangeano dual, P_2^D , é dado por:

$$(P_2^D) \quad \max_{\lambda \geq 0} L_2^{\text{TGP}}(\lambda), \quad (5.23)$$

caso se utilize a relaxação #3, o problema lagrangeano dual, P_3^D , é dado por:

$$(P_3^D) \quad \max_{\substack{\pi \in \mathbb{R} \\ \lambda \geq 0}} L_3^{\text{TGP}}(\pi, \lambda), \quad (5.24)$$

ou, finalmente, caso se utilize a relaxação #4, o problema lagrangeano dual, P_4^D , é dado por:

$$(P_4^D) \quad \max_{\substack{\phi \in \mathbb{R} \\ \lambda \geq 0}} L_4^{\text{TGP}}(\phi, \lambda). \quad (5.25)$$

O método de subgradiente [70] é um dos mais populares para resolução de problemas lagrangeanos duais. Isto porque é de fácil implementação. Os subgradientes são facilmente obtidos a partir das restrições relaxadas e seus resultados são superiores aos demais métodos para alguns problemas em combinatória, principalmente nos casos em que o número de variáveis é muito elevado ou quando ocorre o *gap de dualidade*, sendo necessariamente um método simples. Porém, não existe garantia que os limites inferiores serão crescentes para os multiplicadores de Lagrange gerados por este método. No entanto, na prática bons resultados tem sido obtidos após alguns ajustes, ver [49, 50, 74].

O método de subgradiente pode ser visto como uma adaptação do método de gradiente em que as direções de subida (ou de descida) são geradas a partir dos subgradientes. Isto porque a função objetivo do problema lagrangeano dual (P_1^D , P_2^D , P_3^D ou P_4^D) não é continuamente diferenciável. Em verdade, ele é extremamente simples e pode ser descrito resumidamente da seguinte forma (ver Figura 5.1). Dado um conjunto de multiplicadores de Lagrange, o problema relaxado é solucionado, obtendo-se um limite inferior, e os subgradientes correspondentes a solução relaxada são calculados. Utiliza-se, então, os subgradientes na atualização do conjunto de multiplicadores de Lagrange, visando obter-se um novo limite inferior de valor superior ao anterior. Dessa forma, o método de subgradiente busca gerar limites inferiores crescentes, sendo tal processo repetido até os limites inferior e superior convergirem para um mesmo valor ou a norma do subgradiente ser nula ou, ainda, a existência de um *gap de dualidade* ser detectada.

Como cada uma das relaxações apresentadas possui conjuntos de multiplicadores de Lagrange distintos, elas obviamente terão subgradientes também distintos. Sendo assim, apresenta-se os subgradientes referentes a cada uma delas separadamente.

5.3.1 Relaxação #1

Na relaxação #1, tem-se um conjunto de multiplicadores (π, ϕ, λ) , o qual é utilizado na solução do problema relaxado e, conseqüentemente, na obtenção do limite inferior. Associado à solução relaxada $(\bar{w}_\ell, \bar{f}_\ell^p)$, tem-se os subgradientes $\xi_1^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) = (\xi_{1_{ip}}^{\text{TGP}}(\bar{f}_\ell^p), \xi_{1_i}^{\text{TGP}}(\bar{w}_\ell), \xi_{1_\ell}^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p))$ definidos abaixo:

$$\xi_{1_{ip}}^{\text{TGP}}(\bar{f}_\ell^p) = \sum_{\ell \in A_\ell^+(i)} \bar{f}_\ell^p - \sum_{\ell \in A_\ell^-(i)} \bar{f}_\ell^p - b_i^p, \forall t_p \in P, \forall i \in N_\ell \quad (5.26a)$$

$$\xi_{1_i}^{\text{TGP}}(\bar{w}_\ell) = \sum_{\ell \in A_\ell^+(i)} \bar{w}_\ell - \sum_{\ell \in A_\ell^-(i)} \bar{w}_\ell, \forall i \in N_\ell^o \quad (5.26b)$$

$$\xi_{1_\ell}^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) = \sum_{t_p \in P} \delta_p \bar{f}_\ell^p - \Delta \bar{w}_\ell, \forall \ell \in A_\ell \quad (5.26c)$$

Na k -ésima iteração, um novo conjunto de multiplicadores de Lagrange é obtido da seguinte forma:

$$(\pi, \phi, \lambda)^{k+1} = (\pi, \phi, \lambda)^k + t^k d_1^k(\bar{w}_\ell, \bar{f}_\ell^p) \quad (5.27)$$

em que t^k é um escalar positivo (tamanho do passo) e $d_1^k(\bar{w}_\ell, \bar{f}_\ell^p)$ é uma direção de subida gerada a partir de ξ_1^{TGP} .

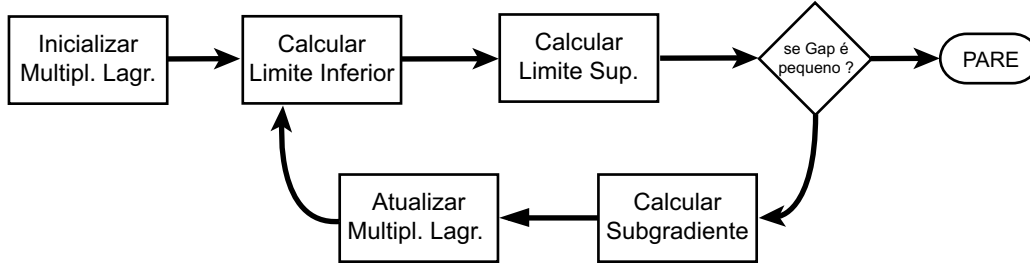


Figura 5.1: Método de subgradientes

5.3.2 Relaxação #2

Na relaxação #2, tem-se um conjunto de multiplicadores $\lambda \geq 0$, o qual é utilizado na solução do problema relaxado e, conseqüentemente, na obtenção do limite inferior. Associado à solução relaxada $(\bar{w}_\ell, \bar{f}_\ell^p)$, tem-se os subgradientes $\xi_2^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) = (\xi_{2_\ell}^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p))$ definidos abaixo:

$$\xi_{2_\ell}^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) = \sum_{t_p \in P} \delta_p \bar{f}_\ell^p - \Delta \bar{w}_\ell, \quad \forall \ell \in A_\ell \quad (5.28)$$

Na k -ésima iteração, um novo conjunto de multiplicadores de Lagrange é obtido da seguinte forma:

$$\lambda^{k+1} = \lambda^k + t^k d_2^k(\bar{w}_\ell, \bar{f}_\ell^p) \quad (5.29)$$

em que t^k é um escalar positivo (tamanho do passo) e $d_2^k(\bar{w}_\ell, \bar{f}_\ell^p)$ é uma direção de subida gerada a partir de ξ_2^{TGP} .

5.3.3 Relaxação #3

Já na relaxação #3, tem-se um conjunto de multiplicadores (π, λ) , o qual é utilizado na solução do problema relaxado e, conseqüentemente, na obtenção do limite inferior. Associado à solução relaxada $(\bar{w}_\ell, \bar{f}_\ell^p)$, tem-se os subgradientes $\xi_3^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) = (\xi_{3_{ip}}^{\text{TGP}}(\bar{f}_\ell^p), \xi_{3_\ell}^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p))$ definidos abaixo:

$$\xi_{3_{ip}}^{\text{TGP}}(\bar{f}_\ell^p) = \sum_{\ell \in A_\ell^+(i)} \bar{f}_\ell^p - \sum_{\ell \in A_\ell^-(i)} \bar{f}_\ell^p - b_i^p, \quad \forall t_p \in P, \forall i \in N_\ell \quad (5.30a)$$

$$\xi_{3_\ell}^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) = \sum_{t_p \in P} \delta_p \bar{f}_\ell^p - \Delta \bar{w}_\ell, \quad \forall \ell \in A_\ell \quad (5.30b)$$

Na k -ésima iteração, um novo conjunto de multiplicadores de Lagrange é obtido da seguinte forma:

$$(\pi, \lambda)^{k+1} = (\pi, \lambda)^k + t^k d_3^k(\bar{w}_\ell, \bar{f}_\ell^p) \quad (5.31)$$

em que t^k é um escalar positivo (tamanho do passo) e $d_3^k(\bar{w}_\ell, \bar{f}_\ell^p)$ é uma direção de subida gerada a partir de ξ_3^{TGP} .

5.3.4 Relaxação #4

Finalmente, na relaxação #4, tem-se um conjunto de multiplicadores (ϕ, λ) , o qual é utilizado na solução do problema relaxado e, consequentemente, na obtenção do limite inferior. Associado à solução relaxada $(\bar{w}_\ell, \bar{f}_\ell^p)$, tem-se os subgradientes $\xi_4^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) = (\xi_{4_i}^{\text{TGP}}(\bar{w}_\ell), \xi_{4_\ell}^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p))$ definidos abaixo:

$$\xi_{4_i}^{\text{TGP}}(\bar{w}_\ell) = \sum_{\ell \in A_\ell^+(i)} \bar{w}_\ell - \sum_{\ell \in A_\ell^-(i)} \bar{w}_\ell, \quad \forall i \in N_\ell^o \quad (5.32a)$$

$$\xi_{4_\ell}^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) = \sum_{t_p \in P} \delta_p \bar{f}_\ell^p - \Delta \bar{w}_\ell, \quad \forall \ell \in A_\ell \quad (5.32b)$$

Na k -ésima iteração, um novo conjunto de multiplicadores de Lagrange é obtido da seguinte forma:

$$(\phi, \lambda)^{k+1} = (\phi, \lambda)^k + t^k d_4^k(\bar{w}_\ell, \bar{f}_\ell^p) \quad (5.33)$$

em que t^k é um escalar positivo (tamanho do passo) e $d_4^k(\bar{w}_\ell, \bar{f}_\ell^p)$ é uma direção de subida gerada a partir de ξ_4^{TGP} .

5.3.5 Direção de Subida

As primeiras versões do método de subgradiente [70, 125] se utilizavam apenas da informação contida no subgradiente da iteração corrente para calcular a direção de subida, isto é, $d^k = \xi^k$, o que equivale dizer que, para as relaxações apresentadas, ter-se-ia:

$$\begin{aligned} d_1^k(\bar{w}_\ell, \bar{f}_\ell^p) &= \xi_1^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) \quad \text{ou} \quad d_2^k(\bar{w}_\ell, \bar{f}_\ell^p) = \xi_2^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) \quad \text{ou} \\ d_3^k(\bar{w}_\ell, \bar{f}_\ell^p) &= \xi_3^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) \quad \text{ou} \quad d_4^k(\bar{w}_\ell, \bar{f}_\ell^p) = \xi_4^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p). \end{aligned}$$

Entretanto, logo se notou que, caso se utilizasse alguma informação da direção de subida da iteração anterior, d^{k-1} , poder-se-ia obter uma melhoria de desempenho. A equação de atualização da direção de subida dada por (5.34) tem se mostrado mais efetiva desde que se faça escolhas razoáveis para o valor de ρ^k .

$$d^k = \xi^k + \rho^k d^{k-1} \quad (5.34)$$

A opção mais simples, denominada *regra de Crowder* [37], utiliza um valor de ρ^k fixo e menor que 1, podendo, ainda, ser formulada como em [74]:

$$d^k = \frac{\xi^k + \rho d^{k-1}}{1 + \rho} \quad (5.35)$$

em que ρ fixo é igual a 0,7.

Uma outra possível abordagem é se utilizar de uma regra mais sofisticada, como a *regra de Camerini–Fratta–Maffioli* [26], dada por:

$$\rho^k = \begin{cases} \frac{-\mu \xi^k d^{k-1}}{\|d^{k-1}\|^2}, & \text{se } \xi^k d^{k-1} < 0 \\ 0, & \text{caso contrário} \end{cases} \quad (5.36)$$

em que μ é um parâmetro a ser estabelecido através de experimentos computacionais (sendo que em [26], os autores indicam que 1,5 é usualmente uma “boa” escolha). O raciocínio por detrás de (5.36) reside no fato de que, ao se escolher adequadamente o valor de μ dentro do intervalo $[0, 2]$, garante-se que a direção d^k é pelo menos tão boa quanto ξ^k .

A necessidade de se atualizar o parâmetro μ pode ser contornada através da adoção da regra de auto-ajuste: $\mu = -\|\xi^k\| \cdot \|d^{k-1}\|/\xi^k d^{k-1}$, que tem por base argumentos geométricos [26]. Dessa forma, obtém-se a seguinte *regra modificada de Camerini–Fratta–Maffioli* [26]:

$$\rho^k = \begin{cases} \frac{\|\xi^k\|}{\|d^{k-1}\|} & , \text{ se } \xi^k d^{k-1} < 0, \\ 0 & , \text{ caso contrário.} \end{cases} \quad (5.37)$$

5.3.6 Tamanho do Passo

A convergência do método de subgradiente, ou melhor, a velocidade com que o mesmo converge para o ótimo está intimamente relacionada com o tamanho do passo. Entre outras formas, o tamanho do passo na k -ésima iteração, t^k , pode ser caracterizado como sugerido por Polyak em [125] da seguinte forma:

$$t^k = \sigma^k \frac{\bar{L} - L^k}{\|\xi^k\|^2} \quad (5.38)$$

em que σ^k é um escalar satisfazendo $0 < \sigma^k \leq 2$, \bar{L} é uma estimativa da solução do problema lagrangeano dual (ou ainda, para o valor do maior limite inferior), e L^k , representa o limite inferior calculado na k -ésima iteração conforme a relaxação utilizada ($L_1^{\text{TGP}}(\pi, \phi, \lambda)$, $L_2^{\text{TGP}}(\lambda)$, $L_3^{\text{TGP}}(\pi, \lambda)$ ou $L_4^{\text{TGP}}(\phi, \lambda)$). Finalmente, $\|\xi^k\|$ é uma norma qualquer, em geral, a norma euclidiana do respectivo subgradiente (ξ_1^{TGP} , ξ_2^{TGP} , ξ_3^{TGP} ou ξ_4^{TGP}) na k -ésima iteração.

Maiores detalhes sobre os aspectos computacionais e propriedades de convergência teórica do método são discutidos em [56, 70]. Segundo a proposta em [70], deve se fazer $\sigma^k = 2$ durante $2n$ iterações, onde n representa o número de variáveis, então divide-se σ^k por dois, após as $2n$ iterações, e seguida, após n iterações, $n/2$, $n/4$, $n/8$, $n/16$, \dots , até se atingir o número mínimo de iterações. Segue-se dividindo σ^k por dois, após cada número mínimo de iterações, até que σ^k se torne suficientemente pequeno. Também pode se definir σ^k como uma função contínua dependente do número de iterações. Nesse caso σ^k pode ser uma função exponencial, conforme o trabalho em [128]. No entanto, no trabalho em [16], bem como em [106, 107, 108, 109, 123, 124], que seguem a proposta de Fisher em [49, 50], σ^k é inicializado como valor igual a 2 e dividido por dois se o melhor limite inferior, L_{max} , não crescer em 30 iterações do método de subgradiente. Semelhante a essa última abordagem, o trabalho em [74], no qual o método de subgradiente é aplicado a um problema de projeto/planejamento de rede, sugere que se faça $\sigma^1 = 1,1$ e, caso não haja melhora do limite inferior durante um número fixo de iterações, σ^k é dividido por dois. Contudo, σ^k deve ser inicializado novamente toda vez que houver melhora da estimativa de solução do problema lagrangeano dual, \bar{L} .

Os valores dos multiplicadores de Lagrange λ obtidos em (5.27), (5.29), (5.31) e (5.33) não são necessariamente não negativos. Para que essa condição seja satisfeita

os multiplicadores devem ser atualizados da seguinte forma:

$$\lambda^{k+1} = \max \{0, \lambda^k + t^k d^k\}. \quad (5.39)$$

Já no tocante à estimativa da solução do problema lagrangeano dual, \bar{L} , bons resultados podem ser obtidos quando o método de subgradiente é utilizado em associação com uma heurística capaz de gerar soluções viáveis, como aquela apresentada na seção 5.2. Dessa forma, poder-se-ia fazer $\bar{L} = \hat{L}_{\min}$, em que \hat{L}_{\min} representa o melhor (isto é, menor) limite superior gerado até o momento. Uma outra sugestão, apresentada por Bazaraa e Sherali em [15] e utilizada em [74], seria fazer:

$$\bar{L} = \frac{\kappa \hat{L}_{\min} + L^k}{2} \quad (5.40)$$

em que κ é um escalar pouco maior que um (por exemplo, $\kappa = 1,05$).

Tal estimativa da solução do problema lagrangeano dual depende da disponibilidade de uma solução viável. Se esse não for o caso, a abordagem mais comum para se obter tal estimativa é multiplicar o melhor limite inferior, L_{\max} , por uma constante $\tau > 0$. Esse método rudimentar poderia ser melhorado pela adoção de outros parâmetros de modo a permitir um ajuste dinâmico da constante τ . Vale ressaltar, ainda, que inúmeros trabalhos na literatura apontam para o fato de que uma estimativa muito precisa (isto é, um limite superior muito justo) nas primeiras iterações do método de subgradiente pode em alguns casos levar a uma deteriorização do desempenho do método.

5.3.7 Algoritmo de Subgradiente

Uma vez descritos os pontos básicos do método de subgradiente, segue-se uma descrição detalhada do algoritmo de subgradiente utilizado na resolução do TGP:

Algoritmo de Subgradiente

- Passo 1** faça $\hat{A}_0 = A_\ell$ e determine uma solução viável e seu valor \hat{L} .
faça $\hat{L}_{\min} = \hat{L}$.
se $\hat{L}_{\min} = +\infty$, então **PARE** (problema é inviável).
- Passo 2** inicialize os multiplicadores de Lagrange.
- Passo 3** calcule o limite inferior L (L_1^{TGP} , L_2^{TGP} , L_3^{TGP} ou L_4^{TGP}).
faça $L_{\max} = L$.
- Passo 4** faça $\hat{A}_0 = \{\ell \in A_\ell \mid \bar{w}_\ell = 1\}$.
calcule uma solução viável e seu valor \hat{L} .
faça $\hat{L}_{\min} = \min\{\hat{L}_{\min}, \hat{L}\}$.
- Passo 5** calcule $Gap = (\hat{L}_{\min} - L_{\max})/\hat{L}_{\min}$.
faça $Iter = 1$, $Non_Increasing = 0$ e $\sigma = 1,1$.
- Passo 6** enquanto ($Gap > \epsilon_1$) e ($\sigma > \epsilon_2$) e ($Iter < Max_Iter$) faça
- Passo 6.1** se $Non_Increasing > Max_Non_Increasing$ então
faça $\sigma = \sigma/2$.
faça $Non_Increasing = 0$.
fimse
- Passo 6.2** calcule os subgradientes ξ (ξ_1^{TGP} , ξ_2^{TGP} , ξ_3^{TGP} ou ξ_4^{TGP}).
se $\|\xi\| < \epsilon_3$, então **PARE**.

- Passo 6.3** calcule o tamanho do passo t .
atualize os multiplicadores de Lagrange.
- Passo 6.4** calcule o limite inferior L (L_1^{TGP} , L_2^{TGP} , L_3^{TGP} ou L_4^{TGP}).
se $L > L_{\max}$ então
faça $\sigma = 1,1$ e $L_{\max} = L$.
faça $\text{Non_Increasing} = 0$.
fimse
- Passo 6.5** faça $\hat{A}_0 = \{\ell \in A_\ell \mid \bar{w}_\ell = 1\}$.
calcule uma solução viável e seu valor \hat{L} .
faça $\hat{L}_{\min} = \min\{\hat{L}_{\min}, \hat{L}\}$.
- Passo 6.6** calcule $\text{Gap} = (\hat{L}_{\min} - L_{\max})/\hat{L}_{\min}$.
faça $\text{Iter} = \text{Iter} + 1$.
fimenquanto

As condições de parada do algoritmo são cinco:

- P1** quando o problema for inviável,
P2 quando o valor do gap é pequeno (isto é, se os limites inferior e superior se equivalem),
P3 quando o valor de σ for reduzido excessivamente,
P4 quando o número máximo de iterações for alcançado, ou
P5 quando a norma do vetor de subgradientes for praticamente nula.

Além disso, no **Passo 6.3**, o cálculo do tamanho do passo utiliza a estimativa dada em (5.40) com $\kappa = 1,05$; enquanto que a atualização dos multiplicadores faz uso das direções de subida dadas por (5.35) com $\rho = 0,7$.

Em resumo, a cada solução do problema lagrangeano dual gerada pelo método de subgradiente, calcula-se o limite inferior correspondente. Tal solução não é necessariamente viável. No entanto, a partir dela pode-se buscar por uma solução viável e, portanto, um limite superior, utilizando-se a heurística descrita anteriormente. O processo iterativo prossegue até que sejam atendidas as condições de otimalidade (ou de parada).

Concluindo, a estratégia para aplicação da relaxação lagrangeana na resolução do **TGP** consiste em encontrar uma solução para o problema lagrangeano dual. Feito isso, calcula-se um limite inferior e, eventualmente, atualiza-se o limite superior até que as condições de otimalidade sejam atendidas.

No entanto, independente dos métodos utilizados, não se garante que uma solução do problema lagrangeano dual possa gerar uma solução ótima para o primal **TGP**. Neste caso tem-se um *gap de dualidade* que pode ser contornado através da ramificação (ou separação) do problema inicial — via algoritmo de *branch-and-bound*.

5.4 Algoritmo de *Branch-and-Bound* para TGP

O algoritmo de *branch-and-bound* foi descrito em detalhes na seção 4.1. Durante o processo de busca no espaço de soluções do problema, um algoritmo de *branch-and-*

bound produz um conjunto de subproblemas a partir do problema original através de sucessivas ramificações (*branching*). Esse conjunto pode ser visto como uma árvore de busca na qual um subproblema é denominado *filho* de um outro subproblema (por sua vez, denominado *pai*) caso o primeiro tenha sido gerado a partir da ramificação (ou separação) do segundo.

Além disso, na árvore de busca, distinguem-se três tipos básicos de nós: inicial, intermediários e terminais (ou finais). O tratamento de cada um desses tipos é geralmente diferenciado e é detalhado em seguida.

5.4.1 Nó Inicial

Corresponde ao problema original. Determina-se uma solução inicial (limite superior) por intermédio de uma heurística. O método de subgradiente é aplicado conforme seção anterior. Se ocorrer a condição de parada **P2** ou **P5**, então a solução ótima para o problema foi encontrada. Caso ocorra **P3** ou **P4**, então deve-se passar a separação (*branching*) do problema (nó) em dois outros. Neste caso, pode-se ainda tentar a aplicação do método de subgradiente para um número fixo de iterações e um valor fixo para σ .

O processo de separação (*branching*) consiste em dividir um problema, ou nó, em um conjunto de subproblemas, ou nós, de tal forma que a união dos conjuntos de soluções viáveis dos subproblemas corresponda ao conjunto de soluções viáveis do problema original. Para o caso específico, um problema de programação inteira envolvendo variáveis de decisão binárias, separa-se um problema em dois outros, fixando para um dado arco $\ell \in A_\ell$ sua variável w_ℓ em 0 em um subproblema e em 1 em outro. A escolha da variável a ser fixada poderá seguir vários critérios, cuja qualidade deve ser testada.

No caso da relaxação #1 pode se fixar a variável $w_{\hat{\ell}}$ associada ao arco $\hat{\ell} \in A_\ell$, tal que $\hat{c}_{\hat{\ell}} = \min_{\ell \in A_\ell} \{\hat{c}_\ell\}$, em que \hat{c}_ℓ representa a contribuição do arco $\ell \in A_\ell$ para a função objetivo do primeiro limite inferior, $L_1^{\text{TGP}}(\pi, \phi, \lambda)$. Analogamente, para a relaxação #2 pode se fixar a variável $w_{\hat{\ell}}$ associada ao arco $\hat{\ell} \in A_\ell$, tal que $\hat{c}_{\hat{\ell}} = \min_{\ell \in A_\ell} \{c_\ell - \Delta\lambda_\ell\}$, em que $c_\ell - \Delta\lambda_\ell$ representa a contribuição do arco $\ell \in A_\ell$ para a função objetivo do segundo limite inferior, $L_2^{\text{TGP}}(\lambda)$. Esse mesmo critério de seleção poderia ser adotado para se fixar uma variável no caso da relaxação #3. Já para a relaxação #4, pode se fixar a variável $w_{\hat{\ell}}$ associada ao arco $\hat{\ell} \in A_\ell$, tal que $\hat{c}_{\hat{\ell}} = \min_{\ell \in A_\ell} \{\tilde{c}_\ell\}$, em que \tilde{c}_ℓ representa a contribuição do arco $\ell \in A_\ell$ para a função objetivo do quarto limite inferior, $L_4^{\text{TGP}}(\phi, \lambda)$.

5.4.2 Nós Intermediários

Os subproblemas resultantes da separação de um problema pai, ou nó pai, representam os nós intermediários da árvore.

A escolha do nó intermediário a ser explorado pode também seguir critérios diversos, entre eles:

- Explorar primeiro o último subproblema gerado. Este critério procura por uma solução o mais rápido possível, de modo a permitir podar um número maior de nós.

- Selecionar o nó com o menor limite inferior, possível candidato a uma solução a baixo custo.

Nestes critérios existe um compromisso em reduzir o espaço de memória ocupada, caminhado pela árvore em profundidade, no primeiro caso; enquanto que no segundo procura-se reduzir o tempo computacional, caminhando na árvore em largura. Pode-se trabalhar com um ou outro ou, até mesmo, os dois intercalados. Na atual implementação optou-se pelo primeiro. Além disso, optou-se por explorar primeiramente o subproblema em que a variável fixada possuía valor igual ao da solução obtida para último limite inferior calculado durante a exploração do nó pai. Dessa forma, em conformidade com o exposto em [74], espera-se acelerar o processo de obtenção de uma solução viável, visto que tal procedimento tende a manter a mesma direção (subgradiente) de busca, ao passo que o subproblema se torna mais simples (devido a fixação de valor das variáveis).

O procedimento de retrocesso (*backtracking*) no caminhamento pela árvore acontece quando $L > \hat{L}_{min}$, quando o subproblema for inviável ou quando o nó for terminal.

De forma a se determinar a viabilidade do subproblema utilizou-se a heurística descrita anteriormente para cálculo do limite superior (uma vez que ela é capaz de produzir soluções de forma bem rápida). Caso a heurística tenha sucesso na obtenção de uma solução viável a partir das facilidades disponíveis (isto é, utilizando apenas aquelas fixadas abertas e as livres), o subproblema certamente é viável e o valor da solução ótima encontrada até o momento pode, eventualmente, ser atualizado (se esse novo limite superior possuir valor menor que o da melhor solução). Caso a heurística venha a falhar, utiliza-se um segundo procedimento de sondagem da viabilidade do subproblema, que consiste na resolução do problema de fluxo de custo mínimo de vários produtos em que os arcos (facilidades) fechados são eliminados. Novamente, se uma solução viável for obtida ela pode ser utilizada na atualização do valor ótimo (isto é, do melhor limite superior encontrado durante a busca). Entretanto, se esse procedimento falhar, não há garantias de que o subproblema seja viável ou não e, portanto, deve se prosseguir na busca.

Sendo assim, essa estratégia para verificação da viabilidade do subproblema não é capaz de determinar se o mesmo é inviável, pois para tanto seria necessário se resolver um problema tanto difícil quanto o problema original. Contudo, ao produzir novos limites superiores e, até mesmo, novas soluções globais para o problema, ela resulta em uma redução do espaço de busca, tendo em vista que muitos subproblemas serão abandonados (*fathomed nodes*), na medida que se verificar que seus limites inferiores superam o valor da melhor solução obtida até um dado momento.

Escolhido o nó e verificada sua viabilidade, a sua exploração consiste em aplicar o método de subgradiente sob condições especiais. Parte-se do conjunto de multiplicadores correspondentes ao maior limite inferior do nó pai. O número de iterações máximo é fixado em 10 e valor de σ é fixado inicialmente como 0,5 (e dividido por 2 se ocorrerem duas interações consecutivas sem que haja aumento do valor do melhor limite inferior). A cada iteração a busca por uma solução viável é realizada utilizando a heurística primal baseada no cálculo de caminhos mínimos.

5.4.3 Nó Terminal

Um nó terminal é aquele em que todas as variáveis de decisão possuem valores fixados em zero ou um. Neste caso, após se verificar que o conjunto de variáveis de decisão atende as restrições (3.10c), basta obter a solução para um problema de fluxo de custo mínimo de vários produtos e retomar a busca a partir do próximo nó (subproblema). Em verdade, é necessário apenas se verificar a viabilidade do problema de fluxo de custo mínimo de vários produtos, pois como os custos estão associados apenas as variáveis de decisão é possível se calcular o valor da solução associada a um nó terminal sem resolver o problema de fluxo de custo mínimo de vários produtos completamente.

5.5 Heurística Lagrangeana para TGP

Entre os objetivos deste trabalho, encontra-se a implementação e teste de uma heurística lagrangeana. Para tanto, o **TGP** foi relaxado (via relaxação lagrangeana) para o cálculo de limites inferiores conforme já descrito na seção 5.1. Implementou-se, então, um método de resolução do problema lagrangeano dual (método de subgradiente) para atualização desses limites como descrito na seção 5.3. Além disso, para obtenção de soluções viáveis a heurística descrita na seção 5.2 também foi implementada.

Tal heurística lagrangeana pode ser vista como uma restrição da implementação do algoritmo de *branch-and-bound* completo, em que apenas o nó inicial (correspondendo ao problema original) é explorado (ver Figura 5.2). Na realidade, o algoritmo apresentado na subseção 5.3.7 corresponde em linhas gerais à heurística lagrangeana implementada. Apesar de tal abordagem não garantir a obtenção de uma solução ótima, os resultados dos testes com essa heurística são muito promissores e são apresentados na seção seguinte. De todo modo, para os casos em que essa heurística falhar (isto é, quando uma solução ótima não for encontrada), torna-se necessário a adoção de um esquema completo de busca no espaço de soluções utilizando-se um algoritmo de *branch-and-bound*.

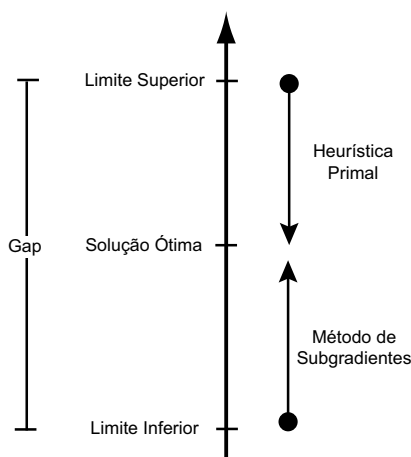


Figura 5.2: Esquema genérico da Heurística Lagrangeana

5.6 Experimentos Computacionais

5.6.1 Resultados da Heurística Lagrangeana

A heurística lagrangeana descrita anteriormente foi implementada utilizando C++. De modo a se resolver o problema de fluxo de custo mínimo, utilizou-se o algoritmo RELAX-IV [19]; enquanto que para o cálculo de caminhos mínimos foi implementado o algoritmo de Dijkstra utilizando *D-Heaps* [146]. Além disso, utilizou-se a biblioteca PPRN [27] para resolver o problema de fluxo de custo mínimo de muitos produtos.

Inicialmente, os valores de todos multiplicadores de Lagrange foram feitos iguais a zero. Além disso, a heurística para cálculo de limite superior foi aplicada a cada iteração¹. Como os testes preliminares apresentaram um melhor desempenho quando todas as três heurísticas para inicialização de \hat{A}_0 (ver seção 5.2) foram utilizadas, optou-se por armazenar todas as soluções viáveis obtidas heurísticamente de modo a se reduzir o tempo total². Por outro lado, para se reduzir o tempo gasto no cálculo do limite inferior, optou-se por utilizar a solução anterior do problema de fluxo de custo mínimo para se obter uma nova, conforme descrito anteriormente.

A heurística lagrangeana foi testada em 150 problemas divididos em 3 conjuntos. O conjunto I contém 90 instâncias do **TGP** para redes em anel com 4, 5, 6, 9, 10 e 16 nós. Para cada anel, padrões de tráfego não uniformes com 40, 80 e 120 produtos foram gerados randomicamente (5 instâncias para cada quantidade de produtos). Todos os produtos possuíam um valor de demanda unitário (isto é, apenas um canal) e o número máximo de produtos transportados em um *container virtual* foi fixado em 63. Já o conjunto II contém 50 instâncias do **TGP** para redes irregulares com 6, 10, 15 e 20 nós. A Figura 5.3 exibe as topologias irregulares usadas para geração das instâncias de teste. As topologia de 6 e 15 nós foram utilizadas em [169], enquanto que a topologia com 20 nós foi usada em [34]. Padrões de tráfego não uniformes foram gerados randomicamente conforme descrito para as redes em anel. Finalmente, o conjunto III contém 10 instâncias baseadas em uma topologia (“real-life”) de rede ótica de transporte Pan-Européia usada em [34]. A Figura 5.4 mostra a topologia da rede.

Para se minimizar tanto o número de *lightpaths* quanto o comprimento dos mesmos (ver discussão no início do capítulo 3 e no final da subseção 3.1.2), os custos estabelecidos na equação (3.7) foram utilizados. Além disso, utilizou-se os seguintes parâmetros³: $Max_Iter = 200$, $Max_Non_Increasing = 30$, $\epsilon_1 = 10^{-3}$ e $\epsilon_2 = \epsilon_3 = 10^{-10}$.

Para o conjunto I, a heurística lagrangeana (utilizando as relaxações #2 e #3) foi capaz de obter a solução ótima de várias instâncias. Os resultados obtidos pela heurística (em um processador Pentium 4 de 1,8 GHz com 512 MB RAM) foram comparados com aqueles obtidos pelo pacote de otimização *CPLEX* [76] – versão 7.0.0 (em uma estação de trabalho SunBlade 100 com um processador RISC de 500 MHz e 1 GB RAM) e são mostrados na Tabela 5.1. A primeira coluna desta tabela apresenta o número de nós na rede original; enquanto que, na segunda coluna,

¹Apesar do cálculo de limites superiores a cada iteração não ser uma prática comum, testes preliminares indicaram que tal abordagem era capaz de produzir soluções viáveis de menor custo.

²O cálculo de limite superior só é realizado uma única vez para um dado subconjunto de arcos.

³Exceto para as redes irregulares de 15 nós com 120 produtos em que $Max_Iter = 400$ e para as redes irregulares de 20 nós com 120 produtos em que $Max_Iter = 600$.

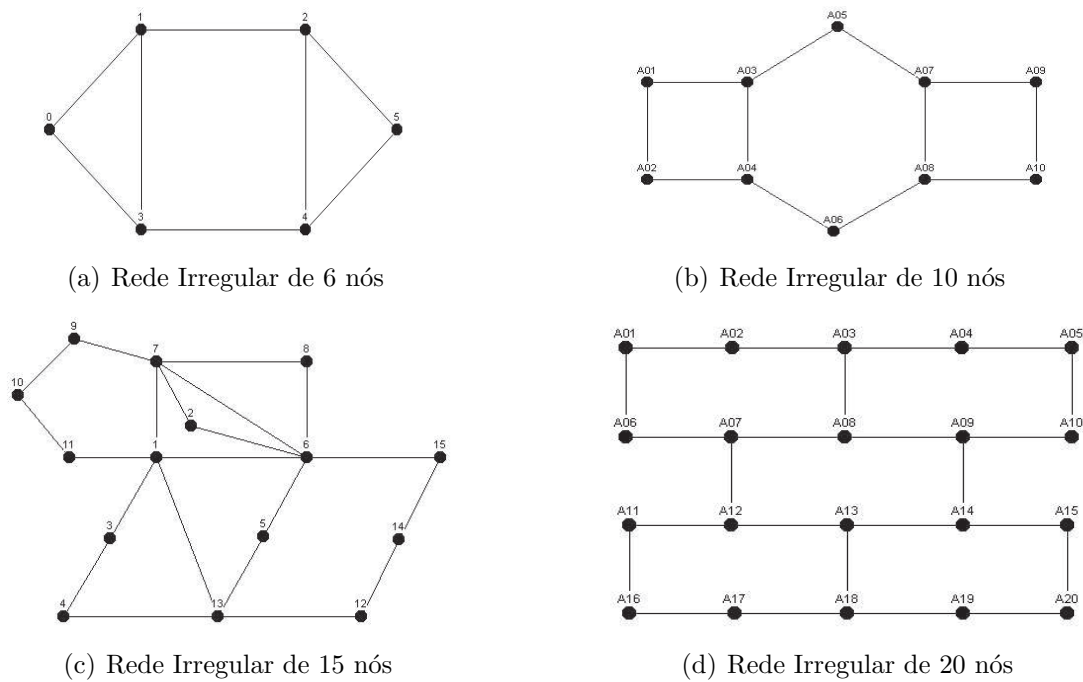


Figura 5.3: Redes com Topologia Irregular

são apresentados o número de nós e o número de arcos da representação estendida associada a cada uma das redes. A terceira coluna exibe o número de produtos de cada subconjunto. O número de variáveis e o número de restrições da formulação de programação matemática são mostrados na quarta e quinta colunas. O número médio de iterações e o tempo médio gasto pelo *CPLEX* para resolver cada subconjunto são apresentados nas duas colunas seguintes, seguidos pelo número médio de iterações, tempo médio e *gap de dualidade* médio obtidos pela heurística lagrangeana utilizando cada uma das 4 relaxações.

Deve-se também mencionar que o *CPLEX* não foi capaz de encontrar uma solução ótima para a maioria das instâncias grandes e sobrecarregadas (como o anel de 16 nós com 120 produtos). Na realidade, mesmo após executar durante 12 horas o

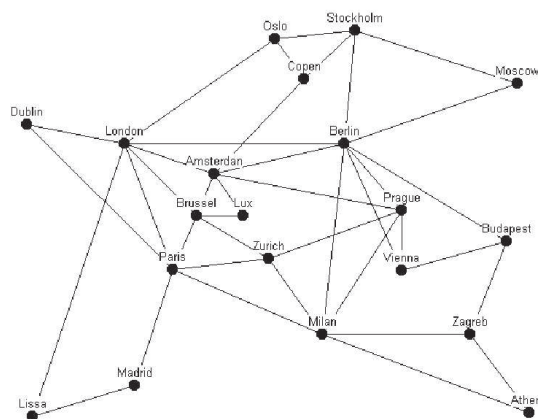


Figura 5.4: Rede Ótica de Transporte Pan-Européia

Tabela 5.1: Resultados da Heurística Lagrangeana para Redes em Anel (valores médios)

V	N A	P	Tam.		CPLEX		Heurística Lagr. - R1			Heurística Lagr. - R2			Heurística Lagr. - R3			Heurística Lagr. - R4		
			Var	Restr	Iter	Tempo (seg)	Iter	Tempo (seg)	Gap ^a (%)	Iter	Tempo (seg)	Gap (%)	Iter	Tempo (seg)	Gap (%)	Iter	Tempo (seg)	Gap (%)
4	20	40	1312	848	1009	4,50	200	0,17	100,00	1	0,00	0,00	1	0,00	0,00	200	0,70	87,06
	32	80	2592	1648	1868	17,52	200	0,32	95,76	1	0,01	0,00	1	0,01	0,00	200	1,55	71,30
	120	120	3872	2448	12664	123,06	200	0,48	52,53	81	1,32	4,14	81	0,42	4,11	200	2,43	60,16
5	25	40	1640	1060	1361	5,69	200	0,23	100,00	1	0,01	0,00	1	0,01	0,00	200	0,98	89,84
	40	80	3240	2060	2469	20,97	200	0,45	77,41	1	0,02	0,00	1	0,01	0,00	200	2,02	77,34
	120	120	4840	3060	28055	242,26	200	0,68	52,66	81	1,77	10,89	81	0,68	3,58	200	3,24	68,37
6	30	40	1968	1272	1763	9,36	200	0,29	100,00	1	0,01	0,00	1	0,01	0,00	200	1,28	91,97
	48	80	3888	2472	3329	39,68	200	0,57	55,66	1	0,03	0,00	1	0,02	0,00	200	2,55	81,05
	120	120	5808	3672	56959	518,03	200	0,86	52,76	121	3,43	3,25	121	0,92	6,07	200	3,94	73,77
9	45	40	2952	1910	3901	21,88	200	0,54	100,00	9	0,13	0,00	50	0,19	0,00	200	2,22	95,68
	72	80	5832	3710	5566	62,11	200	1,04	55,98	1	0,05	0,00	1	0,03	0,00	200	4,32	88,67
	120	120	8712	5510	640755	4743,08	200	1,53	53,00	200	9,29	7,88	200	3,25	6,41	200	6,50	82,84
10	50	40	3280	2120	4545	31,48	200	0,62	100,00	3	0,05	0,00	27	0,13	0,00	200	2,58	94,96
	80	80	6480	4120	6510	85,64	200	1,20	56,19	1	0,06	0,00	1	0,03	0,00	200	4,92	89,60
	120	120	9680	6120	1250768	8956,44	200	1,80	53,06	121	6,74	2,33	121	2,68	0,52	200	7,56	82,77
16	80	40	5248	3392	146905	798,73	200	1,31	100,00	127	3,40	16,91	162	1,24	8,04	200	4,27	97,68
	128	80	10368	6592	13882	181,18	200	2,57	57,19	1	0,14	0,00	1	0,06	0,00	200	8,49	93,89
	120	120	15360	9792	141970	(7,18%)	200	3,88	53,32	160	15,71	4,71	160	5,10	0,76	200	13,68	91,49

^aValores de *gap de dualidade* iguais a 100,00% representam o fato de que os limites inferiores permaneceram iguais a zero ao final da busca

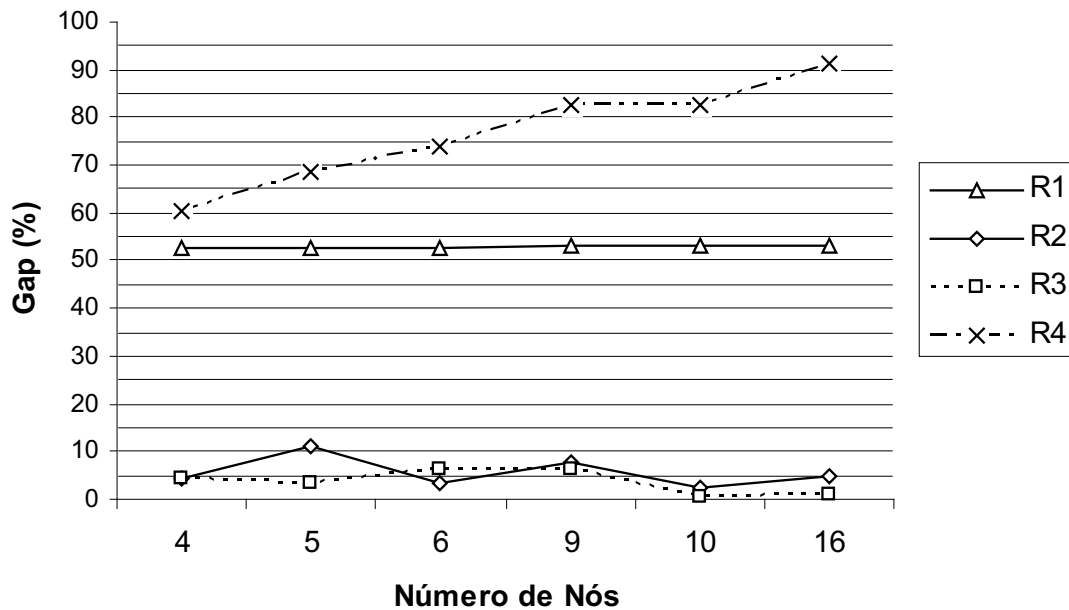


Figura 5.5: Gap Médio das Redes em Anel

CPLEX não obteve nenhuma melhoria significativa depois da primeira hora. Sendo assim, decidiu-se exibir para esses casos os resultados obtidos durante a primeira hora de execução, sendo que o valor entre parênteses representa *gap de dualidade* médio após esse período (isto é, a média dos valores finais de *gap de dualidade*).

Como pode ser facilmente constatado, a heurística lagrangeana independentemente da relaxação utilizada exibe um desempenho bastante estável. Os melhores resultados foram obtidos quando as relaxações #2 e #3 foram utilizadas, o que parece estar associado ao fato dessas relaxações conservarem mais informações sobre a estrutura do problema original. Nesse caso, a heurística lagrangeana conseguiu obter as soluções ótimas de várias instâncias, exceto para instâncias pequenas e sobrecarregadas (como, por exemplo, os anéis de 4, 5 ou 6 nós com 120 produtos) e algumas instâncias grandes com baixo tráfego (por exemplo, o anel de 16 nós com 40 produtos).

O gráfico da Figura 5.5 compara os valores médios do *gap de dualidade* obtidos ao final da execução da heurística lagrangeana para as redes em anel, considerando um tráfego de 120 produtos. O fato dos resultados obtidos pelas duas melhores relaxações (#2 e #3) se alternarem (ver Figura 5.6) pode ser atribuído a natureza heurística do cálculo de limites superiores que depende (conforme descrito anteriormente) da boa qualidade das soluções associadas aos limites inferiores.

A Tabela 5.2 apresenta os resultados para as redes com topologia irregular. Ela possui as mesmas colunas da Tabela 5.1, exceto pela primeira coluna em que o número de arcos da topologia irregular original também é exibido. Novamente, os resultados obtidos são muito satisfatórios, mesmo para grandes instâncias como a rede irregular de 20 nós com 120 produtos. Novamente, para o conjunto II, a heurística lagrangeana (utilizando as relaxações #2 e #3) foi capaz de encontrar a solução

Tabela 5.2: Resultados da Heurística Lagrangeana para Redes Irregulares (valores médios)

V	N	P	Tam.		CPLEX		HL - R1			HL - R2			HL - R3			HL - R4		
			Var	Restr	Iter.	Tempo (seg)	Iter.	Tempo (seg)	Gap ^a (%)	Iter.	Tempo (seg)	Gap (%)	Iter.	Tempo (seg)	Gap (%)	Iter.	Tempo (seg)	Gap (%)
6	38	40	3116	1628	5458	37,24	200	0,51	100,00	1	0,01	0,00	1	0,01	0,00	200	1,92	93,24
8	76	80	6156	3148	10817	158,99	200	0,98	100,00	1	0,03	0,00	1	0,02	0,00	200	3,91	85,53
		120	9196	4668	330777	3317,78	200	1,47	55,66	160	4,97	0,40	160	2,23	0,58	200	5,44	78,42
10	58	40	4428	2476	8358	72,15	200	0,90	100,00	50	0,90	1,84	119	0,82	0,04	200	3,19	97,39
24	108	80	8748	4796	33238	547,33	200	1,75	95,07	21	0,75	0,00	3	0,06	0,00	200	6,35	92,59
		120	13068	7116	430800	5478,31	200	2,64	53,59	132	7,69	0,43	160	3,59	0,95	200	9,07	88,78
15	99	80	19630	8234	108941	2633,58	200	5,27	100,00	48	4,21	0,07	200	6,89	0,61	200	17,75	95,91
21	230	120	27830	12194	108620	(∞)	400	15,85	54,59	271	37,54	0,16	400	20,73	0,92	400	53,88	91,75
20	116	80	17496	9592	141153	(2,19%)	200	5,63	59,80	178	17,90	0,34	200	6,76	1,08	200	17,41	97,13
48	216	120	26136	14232	74665	(∞)	600	25,29	54,10	445	63,62	1,99	600	27,97	3,71	600	91,15	92,66

^aValores de *gap de dualidade* iguais a 100,00% representam o fato de que os limites inferiores permaneceram iguais a zero ao final da busca

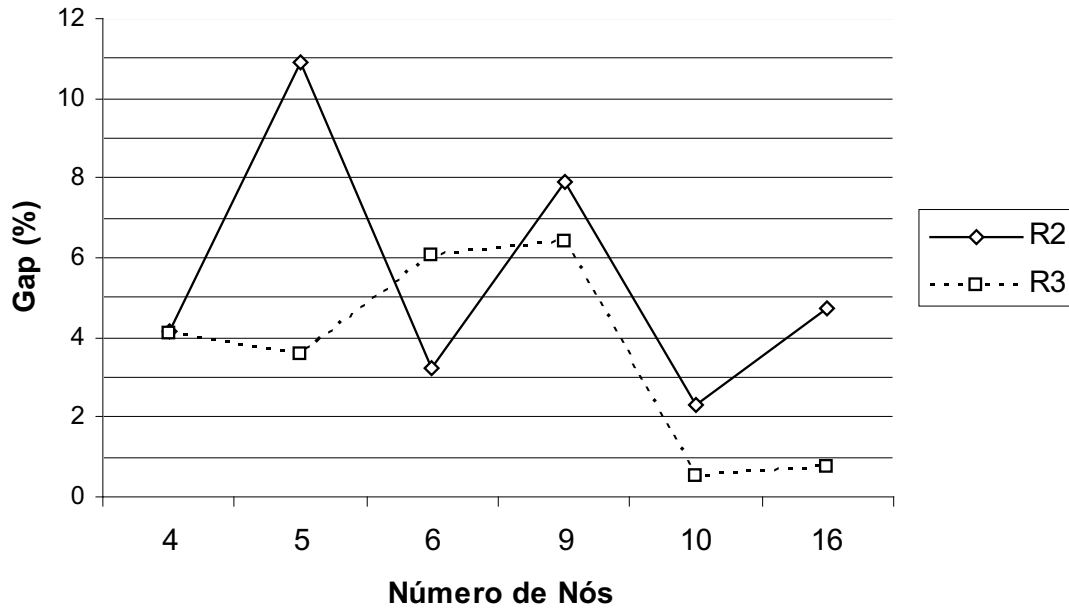


Figura 5.6: Gap Médio para Relaxações #2 e #3 das Redes em Anel

ótima de várias instâncias; enquanto que, para as restantes, ela obteve soluções “quase-ótimas” (isto é, $Gap \leq 1\%$)⁴. As grandes instâncias das redes irregulares (com grande volume de tráfego) parecem ser muito difíceis e o *CPLEX* não conseguiu resolvê-las. De fato, ele não foi capaz de encontrar uma solução viável para a maioria delas durante a primeira hora de processamento ($Gap = \infty$).

Novamente, os resultados apresentados pelas relaxações #2 e #3 são superiores aos demais. O gráfico da Figura 5.7 exhibe os valores médios do *gap de dualidade* obtidos ao final da execução da heurística lagrangeana para as redes irregulares, considerando um tráfego de 120 produtos. Desta vez, contudo, os resultados obtidos pela relaxação #2 superam os da relaxação #3 (ver Figura 5.8). Isso parece estar relacionado ao fato de que, na relaxação #2, o subproblema resolvido envolvendo as variáveis de fluxo é muito mais próximo do problema original que aquele resolvido pela relaxação #3.

Tal comportamento (isto é, o melhor desempenho da relaxação #2) não foi notado nas soluções envolvendo as redes em anel e pode ser explicado da seguinte forma. A principal diferença entre essas duas relaxações está na forma como o subproblema envolvendo as variáveis de fluxo é solucionado. Na relaxação #3, uma vez que se optou pela eliminação (relaxação) das equações de conservação de fluxo, esse subproblema pôde ser resolvido mais facilmente (por inspeção). Na verdade, a partir dos multiplicadores de Lagrange associados às restrições de conservação de fluxo, pode-se obter uma “estimativa” do custo de roteamento do fluxo entre o nó de oferta e o nó de demanda de determinado tráfego. Para uma rede em anel a existência de um único caminho⁵ entre cada par de nós de oferta e de demanda, torna

⁴Para a relaxação #2 foi obtida a solução ótima para 31 das 50 instâncias

⁵Na verdade, existem dois caminhos um no sentido horário e outro no sentido anti-horário

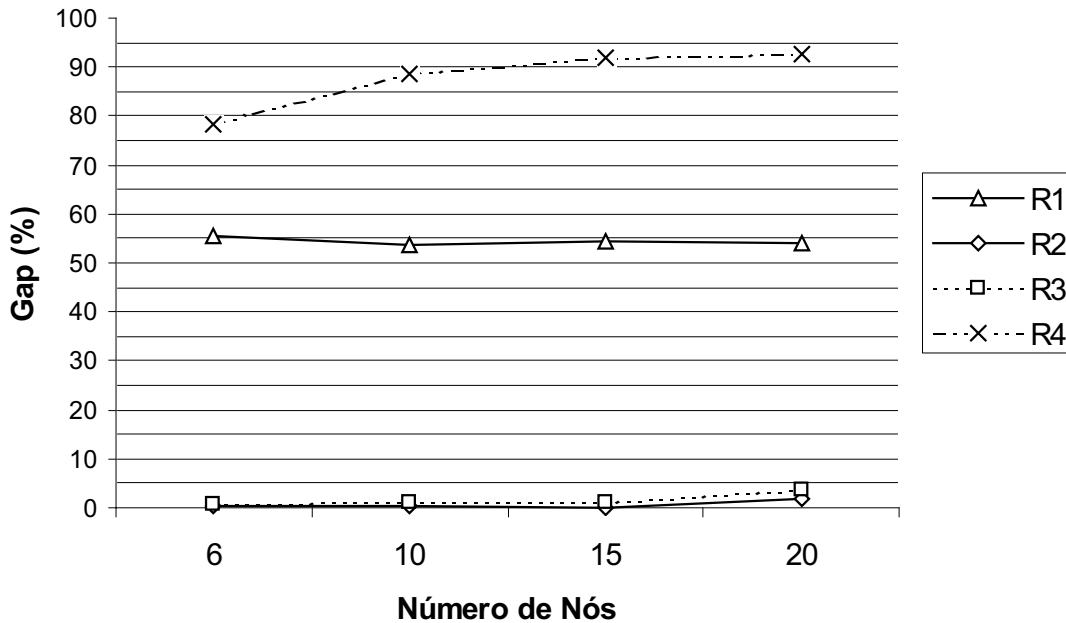


Figura 5.7: Gap Médio das Redes Irregulares

essa estimativa mais próxima do real valor do subproblema de caminho mínimo e, portanto, pouca (ou nenhuma) diferença será percebida. Contudo, na medida que as topologias de rede se tornam mais complexas, como no caso em questão, a equivalência entre esses subproblemas reduz e, portanto, explica as diferenças obtidas nos resultados das relaxações #2 e #3.

A Tabela 5.3 apresenta os resultados detalhados para as instâncias da rede Pan-Européia, além dos valores médios para cada subconjunto. Vale dizer que, para as instâncias com 120 produtos, utilizou-se $Max_Iter = 600$. O *CPLEX* não foi capaz de encontrar nenhuma solução viável durante a primeira hora de processamento e, portanto, seus resultados foram omitidos. Mais uma vez, os resultados são bastante promissores. Apesar dessas instâncias possuírem um grande número de variáveis e restrições, a maioria das soluções é “quase-ótima” e o tempo gasto pela heurística não foi significativo.

Como a heurística lagrangeana obteve bons resultados utilizando as relaxações #2 e #3, realizou-se, em seguida, uma série de experimentos adicionais para se avaliar e comparar os resultados obtidos, caso se modificasse a forma como os multiplicadores de Lagrange foram inicializados. A razão para realização desses testes é procurar determinar se o desempenho da heurística lagrangeana poderia ser melhorado, caso se alterasse o valor inicial dos multiplicadores de Lagrange.

Sendo assim, como a capacidade de um arco é Δ e seu custo c_ℓ , a razão c_ℓ/Δ pode ser entendida como um “custo unitário” para a instalação de capacidade. Foram realizados, então, testes para ambas as relaxações em que o valor inicial dos multiplicadores de Lagrange associados às restrições que limitam a capacidade dos arcos (λ_ℓ^0) foi feito igual a 5% de c_ℓ/Δ . Além disso, também testou-se a relaxação #2 para o caso em que o valor inicial dos mesmos multiplicadores foi feito exatamente

Tabela 5.3: Resultados da Heurística Lagrangeana para a Rede Pan-Européia da Fig. 5.4

V	E	N	A	P	Tam.		Núm. do Probl.	HL - R1			HL - R2			HL - R3			HL - R4					
					Var	Restr		Iter.	Tempo (seg)	Gap ^a (%)	Iter.	Tempo (seg)	Gap (%)	Iter.	Tempo (seg)	Gap (%)	Iter.	Tempo (seg)	Gap (%)			
20	78	176	528	80	42968	14764	1	200	19,29	100,00	200	44,17	0,65	200	22,38	2,03	200	75,98	98,33			
							2	200	19,32	100,00	200	41,09	1,08	200	20,69	1,45	200	72,87	97,72			
							3	200	19,39	100,00	200	58,55	0,33	200	20,83	0,65	200	76,40	98,73			
							4	200	19,31	100,00	200	65,05	0,54	200	20,61	0,81	200	76,99	98,21			
							5	200	19,26	100,00	200	40,11	1,77	200	21,74	1,19	200	71,62	97,20			
							Média	200	19,31	100,00	200	49,794	0,87	200	21,25	1,23	200	74,77	98,04			
							120	63888	21804	1	600	90,00	56,21	600	274,58	0,60	600	92,37	2,87	600	331,27	94,85
										2	600	90,05	100,00	600	295,87	31,06	600	93,71	15,02	600	333,07	94,93
										3	600	90,00	56,15	600	272,01	0,33	600	92,25	14,51	600	324,46	94,72
										4	600	90,07	56,80	600	273,82	0,49	600	94,77	1,56	600	321,83	94,98
				5	600	90,07				56,62	600	258,02	20,24	600	93,15	10,91	600	328,46	94,79			
				Média	600	90,04	65,16	600	274,86	10,54	600	93,25	8,97	600	327,82	94,86						

^aValores de *gap de dualidade* iguais a 100,00% representam o fato de que os limites inferiores permaneceram iguais a zero ao final da busca

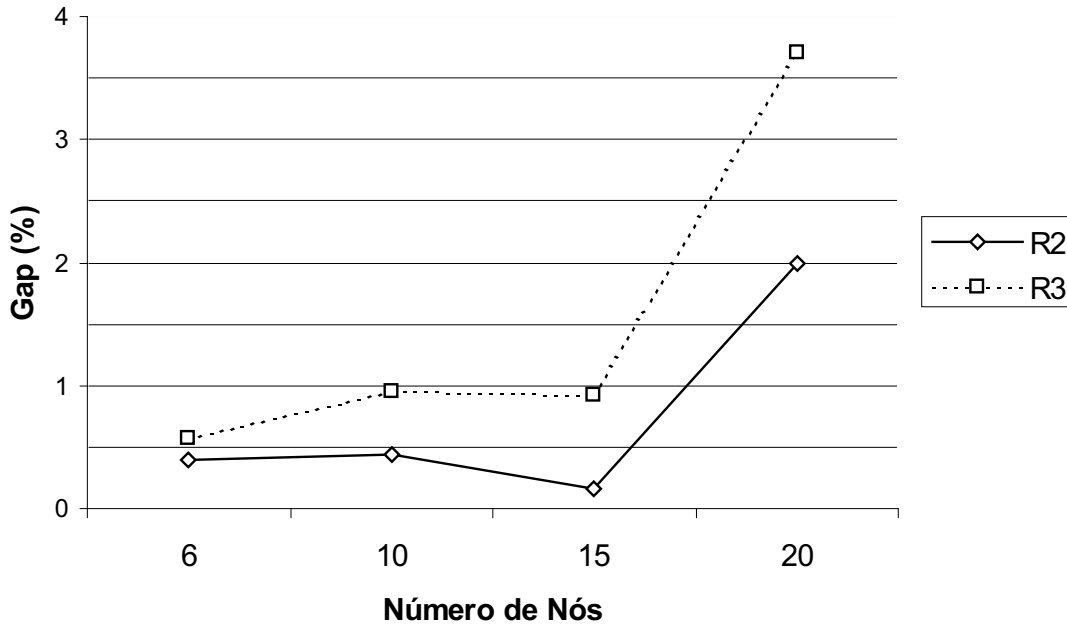


Figura 5.8: Gap Médio para Relaxações #2 e #3 das Redes Irregulares

igual a c_ℓ/Δ^6 . Já o vetor de multiplicadores de Lagrange π (associado às restrições de conservação de fluxo e utilizado na relaxação #3) foi mantido inalterado (isto é, todas as suas componentes permaneceram inicialmente iguais a zero).

Tais experimentos foram realizados com as mesmas instâncias descritas anteriormente (150 problemas divididos em 3 conjuntos) e utilizou-se os seguintes parâmetros⁷: $Max_Iter = 1000$, $\epsilon_1 = 10^{-3}$ e $\epsilon_2 = \epsilon_3 = 10^{-10}$. As Tabelas 5.4 e 5.5 exibem os resultados obtidos nesses experimentos para as redes em anel e irregulares, respectivamente.

Para as redes em anel, constata-se facilmente a superioridade da relaxação #3 em que se inicializou os multiplicadores de Lagrange λ com um valor igual a 5% de c_ℓ/Δ . Neste caso, o valor médio do *gap de dualidade* obtido ao final dos testes é sempre inferior a 4,5% (ver gráficos da Figura 5.9). Observa-se, ainda, que a relaxação #2 em que se inicializou os multiplicadores de Lagrange λ com zero apresentou a maior oscilação em relação as outras duas variações dessa mesma relaxação, ou melhor, um *gap de dualidade*, em média, 4% acima das outras duas para os anéis com 04 e 10 nós. Já para as demais redes em anel as variações da relaxação #2 obtiveram resultados muito próximos (menos de 2% de diferença) e, até mesmo, praticamente iguais.

Já para redes irregulares, os gráficos da Figura 5.10 indicam que a relaxação #3 em que se inicializou os multiplicadores de Lagrange λ com um valor igual a zero apresenta melhores resultados que a outra variação da mesma relaxação (che-

⁶Diferentes valores foram testados inicialmente, porém apenas aqueles que apresentaram diferenças significativas são exibidos adiante.

⁷Exceto para as redes irregulares de 15 nós com 120 produtos em que $Max_Iter = 1500$ e para as redes irregulares de 20 nós com 120 produtos em que $Max_Iter = 2000$.

Tabela 5.4: Resultados das Variações das Relaxações #2 e #3 do **TGP** para Redes em Anel (valores médios)

V	N	A	P	HL - R2			HL - R2 (100%) ^a			HL - R2 (5%) ^b			HL - R3			HL - R3 (5%) ^b		
				Iter	Tempo (seg)	Gap (%)	Iter	Tempo (seg)	Gap (%)	Iter	Tempo (seg)	Gap (%)	Iter	Tempo (seg)	Gap (%)	Iter	Tempo (seg)	Gap (%)
4	20	32	40	1	0,01	0,00	15	0,06	0,03	3	0,02	0,04	1	0,01	0,00	404	0,72	0,10
			80	1	0,01	0,00	145	1,39	0,09	154	1,49	0,09	1	0,01	0,00	503	2,40	0,09
			120	401	6,62	4,14	481	7,99	4,17	499	7,70	8,09	401	3,50	4,11	771	6,30	4,34
5	25	40	40	1	0,01	0,00	21	0,11	0,05	3	0,02	0,02	1	0,01	0,00	452	1,24	0,10
			80	1	0,02	0,00	148	1,81	0,08	120	1,48	0,07	1	0,02	0,00	527	3,84	0,10
			120	401	9,03	3,52	505	11,39	3,63	542	12,09	3,56	401	3,56	3,58	751	9,46	3,64
6	30	48	40	1	0,01	0,00	12	0,08	0,02	3	0,02	0,03	1	0,01	0,00	462	1,23	0,10
			80	1	0,03	0,00	116	1,76	0,08	119	1,85	0,06	1	0,02	0,00	503	3,30	0,10
			120	601	16,81	3,15	670	18,89	3,17	672	19,01	3,17	601	7,34	6,07	834	8,82	0,59
9	45	72	40	10	0,13	0,00	18	0,23	0,03	5	0,07	0,06	50	0,20	0,00	528	2,71	0,10
			80	1	0,05	0,00	134	3,51	0,08	187	4,98	0,09	1	0,03	0,00	584	7,32	0,10
			120	1000	47,68	7,96	1000	47,32	6,22	1000	48,28	6,41	1000	23,71	4,74	1000	21,69	1,08
10	50	80	40	3	0,05	0,00	11	0,17	0,02	4	0,07	0,03	27	0,14	0,00	534	2,65	0,10
			80	1	0,06	0,00	99	3,01	0,06	203	6,29	0,08	1	0,03	0,00	573	6,69	0,10
			120	601	33,59	2,33	661	36,55	0,55	685	39,45	5,44	601	12,00	0,52	839	14,83	1,82
16	80	128	40	605	18,98	16,91	609	19,09	14,96	608	20,66	16,18	642	5,11	8,04	935	6,70	6,14
			80	1	0,13	0,00	96	6,00	0,07	199	12,37	0,09	1	0,06	0,00	546	11,75	0,10
			120	800	86,96	3,73	835	90,04	5,94	850	92,20	4,60	800	30,46	0,76	931	32,47	0,88

^aResultados obtidos quando os multiplicadores de Lagrange λ foram inicializados com o valor c_ℓ/Δ ^bResultados obtidos quando os multiplicadores de Lagrange λ foram inicializados com 5% do valor c_ℓ/Δ

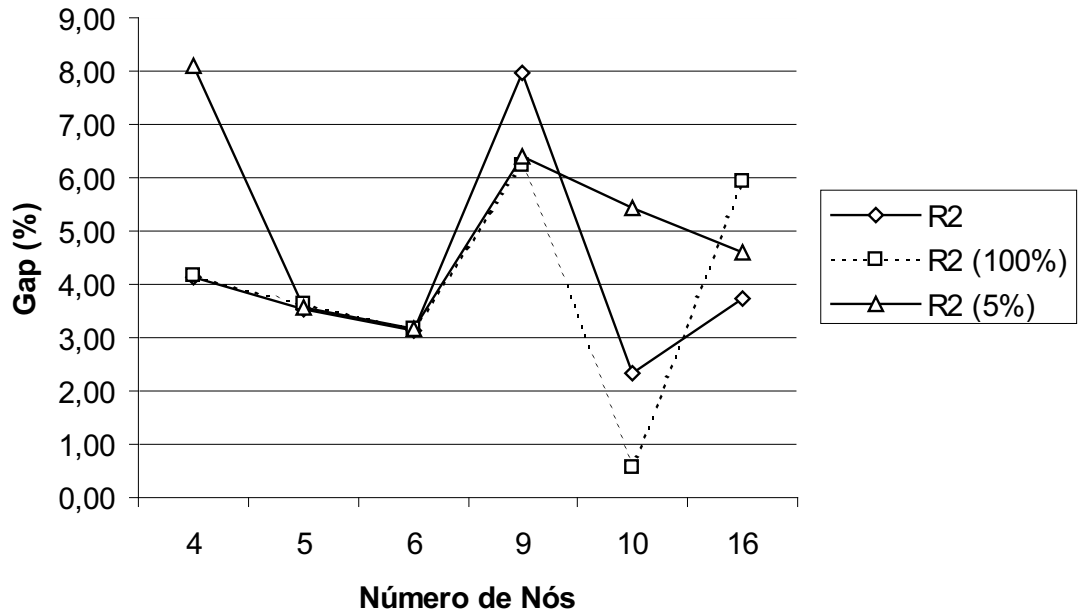
Tabela 5.5: Resultados das Variações das Relaxações #2 e #3 do **TGP** para Redes Irregulares (valores médios)

V	E	N	A	P	HL - R2			HL - R2 (100%) ^a			HL - R2 (5%) ^b			HL - R3			HL - R3 (5%) ^b		
					Iter	Tempo (seg)	Gap (%)	Iter	Tempo (seg)	Gap (%)	Iter	Tempo (seg)	Gap (%)	Iter	Tempo (seg)	Gap (%)	Iter	Tempo (seg)	Gap (%)
6	8	38	76	40	1	0,01	0,00	19	0,16	0,05	8	0,08	0,07	1	0,01	0,00	555	1,88	0,10
				80	1	0,03	0,00	150	2,65	0,08	123	2,28	0,09	1	0,02	0,00	583	4,33	0,10
				120	800	23,51	0,40	866	25,70	0,27	836	25,38	0,16	800	13,00	0,25	906	11,36	0,74
10	24	58	108	40	358	5,71	0,00	23	0,38	0,06	237	4,17	3,35	129	0,95	0,00	558	3,12	0,10
				80	21	0,70	0,00	123	4,08	0,09	134	4,53	0,08	3	0,06	0,00	560	6,34	0,10
				120	618	34,88	0,52	682	38,19	0,38	671	38,90	0,25	801	20,35	0,39	911	17,12	1,04
15	21	99	230	80	43	3,54	0,07	139	10,86	0,09	114	9,39	0,10	615	23,04	0,29	1000	29,07	0,20
				120	415	50,16	0,25	692	82,87	0,10	553	70,46	0,21	1464	83,21	0,59	1285	54,58	5,38
20	48	116	216	80	476	43,50	0,12	525	47,08	0,35	407	40,16	0,11	1000	35,65	0,91	1000	29,61	0,47
				120	585	78,34	0,08	520	72,59	0,09	750	107,39	0,07	1136	60,67	3,71	2000	87,53	1,14
20 ^c	78	176	528	80	1000	198,58	1,02	1000	201,84	0,34	1000	194,98	0,41	1000	119,65	0,62	1000	100,04	0,81
				120	1550	558,18	8,82	1670	610,21	4,88	1919	633,41	0,54	1556	273,99	2,87	1741	263,76	5,72

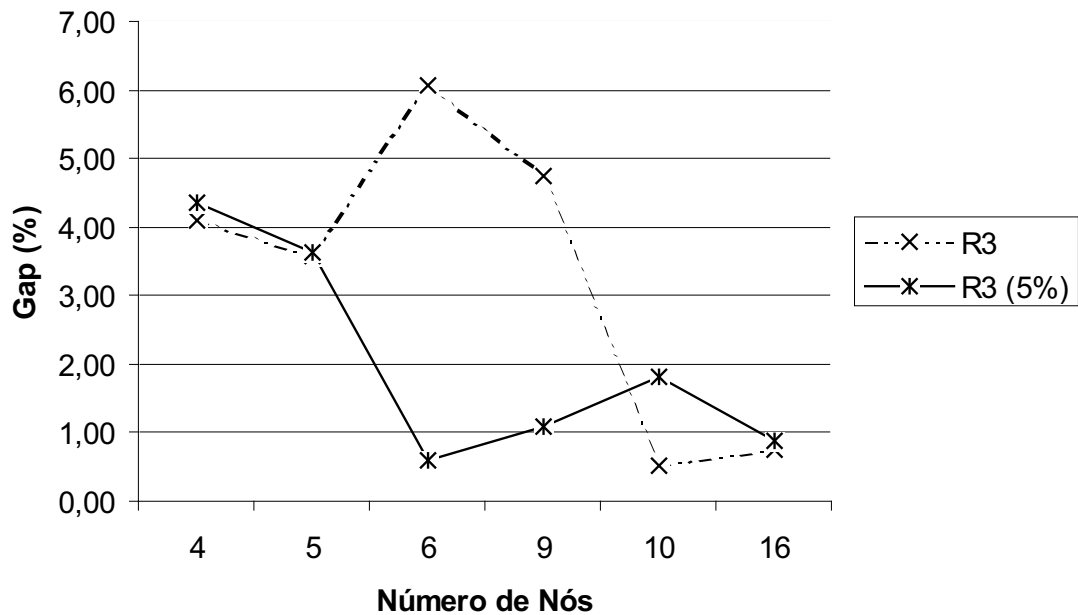
^aResultados obtidos quando os multiplicadores de Lagrange λ foram inicializados com o valor c_ℓ/Δ

^bResultados obtidos quando os multiplicadores de Lagrange λ foram inicializados com 5% do valor c_ℓ/Δ

^cEsses resultados correspondem a rede Pan-Européia

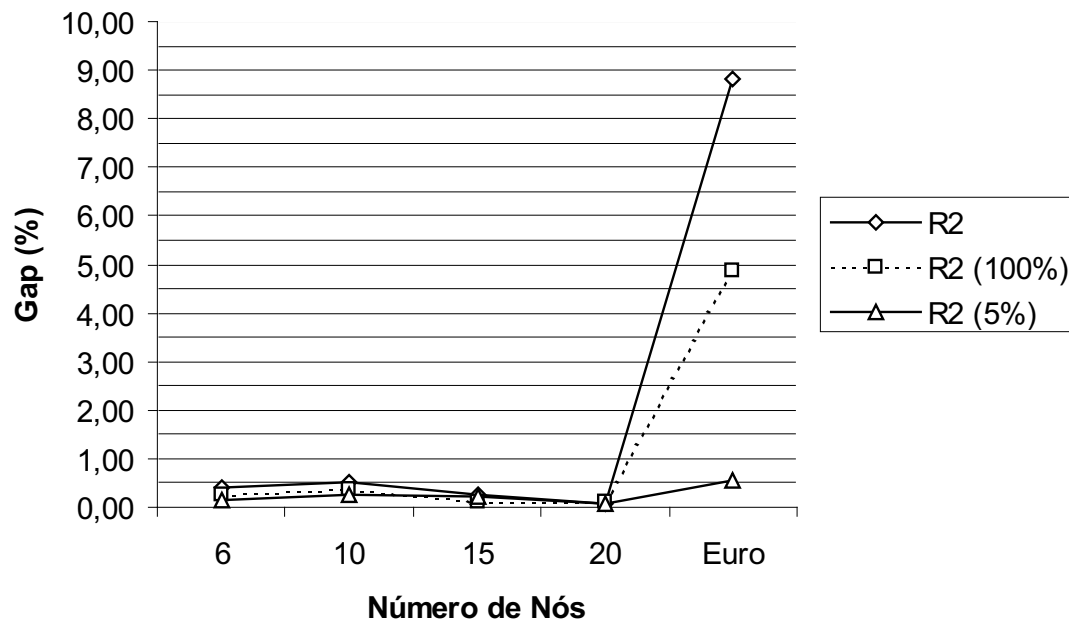


(a) Variações da Relaxação #2

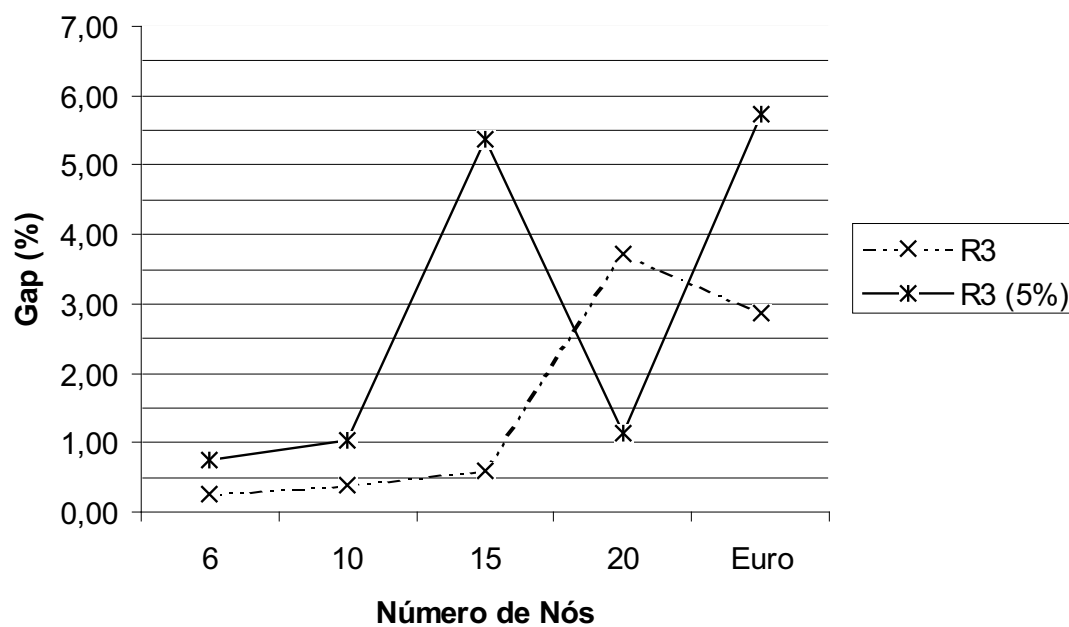


(b) Variações da Relaxação #3

Figura 5.9: Gap Médio das Variações das Relaxações para Redes em Anel



(a) Variações da Relaxação #2



(b) Variações da Relaxação #3

Figura 5.10: Gap Médio das Variações das Relaxações para Redes Irregulares

gam a apresentar uma diferença de aproximadamente 5% em relação a outra para as redes com 15 nós). Além disso, os resultados obtidos parecem indicar que as variações da relaxação #2 são praticamente equivalentes entre si quando aplicadas às redes irregulares. Dessa forma, de modo a diferenciá-las acrescentou-se a Tabela 5.5 e, conseqüentemente, ao gráfico da Figura 5.10 os resultados obtidos para rede Pan-Européia (que também apresenta uma topologia irregular)⁸. Como pode ser facilmente observado, o desempenho da relaxação #2 em que se inicializou os multiplicadores de Lagrange λ com um valor igual a 5% de c_ℓ/Δ parece ser mais consistente (apresentando menores oscilações) que as outras duas variações.

Como futuramente as redes tendem a apresentar topologias irregulares, a escolha mais sensata parece ser utilizar ou a relaxação #2 com os multiplicadores de Lagrange λ inicializados com um valor igual a 5% de c_ℓ/Δ ou a relaxação #3 com todos os multiplicadores (λ e π) inicializados com zero. Além disso, a utilização das relaxações #1 e #4 na implementação de um procedimento de busca (algoritmo de *branch-and-bound*) não parece promissora, uma vez que ambas apresentaram valores médios muito elevados para o *gap de dualidade*.

5.6.2 Resultados do Algoritmo de *Branch-and-Bound*

O algoritmo de *branch-and-bound* descrito anteriormente foi implementado em C++ e testado na mesma plataforma que a heurística lagrangeana. Tomando-se como base os resultados da análise realizada na subseção anterior, optou-se pela implementação de apenas duas versões do algoritmo de *branch-and-bound* uma utilizando a relaxação #2 (com os multiplicadores de Lagrange λ inicializados com um valor igual a 5% de c_ℓ/Δ) e outra utilizando a relaxação #3 (com todos os multiplicadores inicializados com zero). Essas versões foram denominadas **BBR2** e **BBR3**, respectivamente.

Novamente, os experimentos foram realizados com as mesmas instâncias descritas anteriormente (150 problemas divididos em 3 conjuntos), sendo que estabeleceu-se um limite máximo de tempo de 1800 segundos. Cada subconjunto de instâncias é identificado através do seguinte padrão: TX-P, em que T identifica o tipo de topologia da rede (para redes em anel T = C, já para redes irregulares T = M ou N), X indica o número de nós da rede (X = 04, 05, 06, 09, 10, 16 para redes em anel ou X = 06, 10, 15, 20 para redes irregulares) e P indica o número de produtos (P = 40, 80, 120).

As Tabelas 5.6 e 5.7 exibem os resultados médios para as redes em anel. A primeira coluna dessas tabelas identifica o subconjunto de instâncias, enquanto que as três colunas seguintes exibem os valores médios para o número de iterações, tempo gasto (em segundos) e *gap de dualidade* obtidos ao final da exploração do nó inicial. Já as próximas cinco colunas apresentam informações sobre o processo de busca. O número médio de iterações, o tempo médio gasto na busca e o número médio de nós (subproblemas) gerados durante a busca são exibidos na quinta, sexta e sétima colunas, respectivamente. A oitava coluna apresenta o número médio de nós (subproblemas) da árvore de busca efetivamente explorados, enquanto que a nona coluna exhibe o número médio de nós (subproblemas) em que ocorreu um retrocesso (*backtracking*) durante a busca (isto é, sua exploração não foi necessária), conforme

⁸Nestes testes utilizou-se *Max_Iter* = 2000

Tabela 5.6: Resultados do Algoritmo de *Branch-and-Bound* **BBR2** para Redes em Anel (valores médios)

Instância	Nó Inicial			Árvore de Busca					ADMs	LPs	WLs
	Iter	Tempo (seg)	Gap (%)	Iter	Tempo (seg)	Nós Gerados	Nós Expl.	Nós Podados			
C04-40	3	0,02	0,04	3	0,02	1	1	0	4,0	4,0	8,0
C04-80	154	1,51	0,09	154	1,51	1	1	0	4,0	4,0	8,0
C04-120	499	8,40	8,09	1099	16,13	120	62	58	4,4	4,2	9,2
C05-40	3	0,02	0,02	3	0,02	1	1	0	5,0	5,0	10,0
C05-80	120	1,52	0,07	120	1,52	1	1	0	5,0	5,0	10,0
C05-120	542	12,50	3,56	2484	48,72	385	200	184	5,4	5,2	11,6
C06-40	3	0,03	0,04	3	0,03	1	1	0	6,0	6,0	12,0
C06-80	119	1,87	0,06	119	1,87	1	1	0	6,0	6,0	12,0
C06-120	672	19,36	3,17	1875	29,67	216	147	69	6,0	6,0	16,8
C09-40	5	0,08	0,05	5	0,08	1	1	0	9,0	9,0	18,0
C09-80	187	4,93	0,10	187	4,94	1	1	0	9,0	9,0	18,0
C09-120	1000	48,44	2,83	22459	171,83	3630	2636	994	9,0	9,0	32,0
C10-40	4	0,08	0,03	4	0,08	1	1	0	10,0	10,0	20,0
C10-80	203	6,19	0,09	203	6,20	1	1	0	10,0	10,0	20,0
C10-120	679	38,22	0,55	28223	154,48	4690	3339	1352	10,0	10,0	29,6
C16-40	614	19,58	16,20	24940	373,27	4867	2434	2424	18,0	17,2	35,6
C16-80	199	12,13	0,09	199	12,14	1	1	0	16,0	16,0	32,0
C16-120	850	93,28	1,84	137137	1445,06	24459	15219	9221	16,4	16,2	52,0

Tabela 5.7: Resultados do Algoritmo de *Branch-and-Bound* **BBR3** para Redes em Anel (valores médios)

Instância	Nó Inicial			Árvore de Busca					ADMs	LPs	WLs
	Iter	Tempo (seg)	Gap (%)	Iter	Tempo (seg)	Nós Gerados	Nós Expl.	Nós Podados			
C04-40	1	0,01	0,00	1	0,01	1	1	0	4,0	4,0	8,0
C40-80	1	0,02	0,00	1	0,02	1	1	0	4,0	4,0	8,0
C04-120	401	2,97	10,11	2803	18,68	482	249	234	4,2	4,2	9,6
C05-40	1	0,02	0,00	1	0,02	1	1	0	5,0	5,0	10,0
C05-80	1	0,02	0,00	1	0,02	1	1	0	5,0	5,0	10,0
C05-120	401	4,23	3,46	7063	86,36	1334	679	656	5,4	5,2	11,6
C06-40	1	0,01	0,00	1	0,01	1	1	0	6,0	6,0	12,0
C06-80	1	0,03	0,00	1	0,03	1	1	0	6,0	6,0	12,0
C06-120	601	7,63	3,23	28364	243,92	5555	2859	2695	6,0	6,0	16,8
C09-40	57	0,25	0,00	57	0,25	1	1	0	9,0	9,0	18,0
C09-80	1	0,03	0,00	1	0,04	1	1	0	9,0	9,0	18,0
C09-120	1000	21,31	0,84	138308	1800,06	27465	14052	13384	9,0	9,0	32,0
C10-40	9	0,06	0,00	9	0,06	1	1	0	10,0	10,0	20,0
C10-80	1	0,04	0,00	1	0,04	1	1	0	10,0	10,0	20,0
C10-120	601	10,49	0,52	72121	1080,12	14306	7338	6947	10,0	10,0	29,6
C16-40	641	5,53	8,27	59439	766,18	11762	5888	5862	17,2	16,4	40,4
C16-80	1	0,07	0,00	1	0,07	1	1	0	16,0	16,0	32,0
C16-120	800	29,72	1,90	37720	1440,31	7387	3729	3611	16,2	16,2	54,4

descrito na subseção 5.4.2. Por fim, as três últimas colunas, apresentam os valores médios para o número de multiplexadores *add-drop*, para o número de *lightpaths* e para o número de *wavelength-links*⁹ obtidos pela melhor solução encontrada durante o processo de busca.

Ao se analisar as informações sobre a exploração do nó inicial, é fácil perceber que **BBR3** supera **BBR2** para diversos subconjuntos de instâncias. Para esses subconjuntos, **BBR3** realiza um número menor de iterações consumindo, consequentemente, muito menos tempo. Contudo ela é capaz de obter um *gap de dualidade* menor. Essa superioridade de **BBR3** na exploração do nó inicial já era esperada, tendo em vista os resultados apresentados na subseção anterior. Porém, quando se compara as informações relativas ao procedimento de busca propriamente dito, os resultados encontrados indicam que, para instâncias mais “difíceis”, **BBR3** necessita explorar um número maior de nós. Apesar das dificuldades iniciais (isto é, na exploração do nó inicial), **BBR2** conseguiu limitar sua busca a uma pequena parcela dos subproblemas do espaço de soluções, provavelmente devido ao fato de seus limites inferiores possuírem mais informações sobre a estrutura do problema original que os gerados por **BBR3** (em que as restrições de conservação de fluxo foram relaxadas).

Já as Tabelas 5.8 e 5.9 exibem os resultados médios para as redes irregulares e apresentam as mesmas colunas que as tabelas anteriores – Tabelas 5.6 e 5.7. Neste caso, ao se comparar os resultados obtidos por **BBR2** e **BBR3**, observa-se claramente que, na medida que o problema de torna mais difícil (devido ao aumento do número de nós e arcos da rede), **BBR2** apresenta resultados claramente superiores (o que de fato já era esperado, em virtude da análise realizada na subseção anterior).

Essas dificuldades encontradas na limitação da busca se refletem diretamente na quantidade de instâncias para as quais **BBR3** conseguiu encontrar a solução ótima (dentro do limite de tempo estabelecido). **BBR3** conseguiu encontrar e garantir a otimalidade da solução (através da exaustão do processo de busca) de apenas 102 das 150 instâncias de teste (68%), enquanto que **BBR2** obteve a solução ótima para 132 instâncias (88%).

Tais resultados parecem pouco animadores, contudo um exame mais detalhado das soluções obtidas demonstram que, na maioria dos casos, a solução ótima (ou “quase-ótima”) foi obtida na exploração do nó inicial (apenas 7 das soluções encontradas por **BBR2** ocorreram durante a exploração de outro nó da árvore, enquanto que para **BBR3** esse número é de apenas 6). As Tabelas 5.10 e 5.11 exibem o resultado da busca para cada uma das 10 instâncias da rede Pan-Européia e servem para ilustrar esse fato (apenas para uma das instâncias a melhor solução viável foi obtida durante a exploração de um nó diferente do inicial).

Em verdade, se examinarmos os valores médios do *gap de dualidade* obtidos ao final da exploração do nó inicial por ambas as implementações, verifica-se claramente que as soluções encontradas até então devem ser ótimas ou “quase-ótimas”. E, portanto, o processo de busca muito provavelmente não será capaz de melhorar tal valor, restando a ele somente a difícil tarefa de provar a otimalidade da solução

⁹Um *wavelength-link* representa a alocação de um comprimento de onda em um arco da rede. Portanto o número de “saltos” (*hops*) de um *lightpath*, na topologia virtual, é igual ao número de *wavelength-links* associados a ele, ao passo que o número médio de “saltos” por *lightpath* é igual a razão entre o número total de *wavelength-links* e o número total de *lightpaths*.

Tabela 5.8: Resultados do Algoritmo de *Branch-and-Bound* **BBR2** para Redes Irregulares (valores médios)

Instância	Nó Inicial			Árvore de Busca					ADMs	LPs	WLs
	Iter	Tempo (seg)	Gap (%)	Iter	Tempo (seg)	Nós Gerados	Nós Expl.	Nós Podados			
M06-40	8	0,08	0,07	8	0,08	1	1	0	6,0	6,0	12,0
M06-80	123	2,25	0,09	123	2,25	1	1	0	6,0	6,0	12,0
M06-120	836	25,48	0,16	1446	27,32	99	77	21	6,0	6,0	13,6
M10-40	214	3,53	3,36	279	4,57	14	8	6	10,0	10,0	20,0
M10-80	134	4,45	0,09	134	4,45	1	1	0	10,0	10,0	20,0
M10-120	670	38,17	0,25	16530	122,31	2765	1945	820	10,0	10,0	24,0
M15-80	109	8,68	0,10	109	8,69	1	1	0	15,0	15,0	34,0
M15-120	421	53,03	0,13	29513	387,12	5291	3345	1930	15,0	15,0	35,6
N20-80	445	41,77	0,16	47487	383,03	9410	4705	4688	20,0	20,0	46,4
N20-120	695	97,03	0,21	35533	562,17	6498	3792	2692	20,0	20,0	48,8

Tabela 5.9: Resultados do Algoritmo de *Branch-and-Bound* **BBR3** para Redes Irregulares (valores médios)

Instância	Nó Inicial			Árvore de Busca					ADMs	LPs	WLs
	Iter	Tempo (seg)	Gap (%)	Iter	Tempo (seg)	Nós Gerados	Nós Expl.	Nós Podados			
M06-40	1	0,01	0,00	1	0,02	1	1	0	6,0	6,0	12,0
M06-80	1	0,03	0,00	1	0,03	1	1	0	6,0	6,0	12,0
M06-120	800	13,93	0,32	119671	1440,03	23385	12193	11161	6,0	6,0	14,8
M10-40	173	1,29	0,00	173	1,29	1	1	0	10,0	10,0	20,0
M10-80	3	0,07	0,00	3	0,07	1	1	0	10,0	10,0	20,0
M10-120	801	19,34	0,52	60916	1440,09	11967	6125	5795	10,0	10,0	29,6
M15-80	668	25,26	0,23	22297	730,43	4314	2171	2090	15,0	15,0	38,4
M15-120	1000	57,17	0,62	36407	1800,33	7003	3622	3246	15,0	15,0	49,2
N20-80	1000	37,51	0,57	33900	1800,18	6574	3300	3168	20,0	20,0	60,8
N20-120	1000	56,97	2,84	32972	1800,33	6397	3203	3096	20,4	20,4	80,0

encontrada. Isso também leva a crer que, em vista da dificuldade em se provar a otimalidade da solução, “boas” soluções poderiam ser obtidas a partir da restrição do processo de resolução do **TGP** apenas à exploração do nó inicial (isto é, a utilização da heurística lagrangeana) ou, ainda, da aplicação do procedimento de busca de forma limitada (por tempo ou por número de nós).

Por fim, os resultados obtidos para a quantidade de multiplexadores de *add-drop* – ADMs, juntamente com o número de *lightpaths* – LPs (ambos exibidos na antepenúltima e penúltima colunas das tabelas), aliado ao fato de se utilizar apenas um comprimento de onda em todos os testes realizados, indicam a adoção de uma topologia de rede em anel para a rede virtual. Esse fato é examinado mais detalhadamente adiante.

5.7 Comparação com Outras Abordagens

5.7.1 Avaliação Preliminar

A comparação direta dos resultados obtidos com aqueles produzidos por outras abordagens descritas na literatura encontra três grandes dificuldades. A primeira delas está relacionada com a quase completa inexistência de informações detalhadas sobre as instâncias utilizadas nos testes publicados. Em sua maioria, os resultados apresentados na literatura foram obtidos a partir de instâncias geradas aleatoriamente e, mesmo que se utilize os parâmetros descritos em tais artigos, é improvável que se consiga reproduzir fielmente as instâncias que foram utilizadas.

Uma segunda dificuldade se encontra relacionada às restrições e/ou hipóteses “simplificadoras” adotadas na maioria dos trabalhos da literatura. Entre tais restrições, destacam-se várias abordagens que solucionam apenas o **TGP-SH** ou aquelas em que o **TGP-MH** é resolvido considerando a existência de um único *hub* – elemento da rede com capacidade de realizar *grooming* (ver subseção 2.3.2).

Por fim, o tipo de abordagem usado na resolução do problema representa uma última dificuldade a ser considerada. Quase toda a literatura sobre o problema em questão descreve abordagens heurísticas para o mesmo. Tais algoritmos heurísticos lançam mão principalmente de abordagens gulosas, de meta-heurísticas (*simulated annealing*) e de algoritmos genéticos. Isto torna a comparação de tais resultados com aqueles produzidos pelas abordagens propostas neste trabalho uma tarefa árdua, senão impossível.

Apesar das considerações acima, esta subseção procura realizar uma comparação preliminar entre a heurística lagrangeana descrita na seção 5.5 e aquelas apresentadas

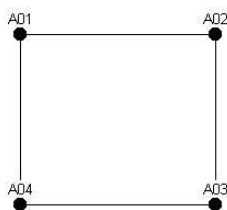


Figura 5.11: Rede em Anel com 4 Nós

Tabela 5.10: Resultados do Algoritmo de *Branch-and-Bound* **BBR2** para Rede Pan-Européia

Instância	Nó Inicial					Árvore de Busca				ADMs	LPs	WLs
	Iter	Tempo (seg)	Limite Inferior	Limite Superior	Gap (%)	Iter	Tempo (seg)	Nós da Árvore Gerados(Expl./Podados)	Solução (Nó da Solução) ^a			
Euro-80-1	1000	202,15	40259,74	40392,00	0,33	106779	1801,07	20989 (10914/9659)	40392,00 (1)*	20,0	20,0	52,0
Euro-80-2	1000	196,54	40259,58	40348,00	0,22	102307	1800,02	19435 (11713/7477)	40348,00 (1)*	20,0	20,0	48,0
Euro-80-3	1000	189,34	40259,71	40634,00	0,92	103067	1800,04	20121 (10456/9349)	40634,00 (1)*	20,0	20,0	74,0
Euro-80-4	1000	227,54	40259,33	40436,00	0,44	103230	1800,10	20529 (10264/9926)	40436,00 (1)*	20,0	20,0	56,0
Euro-80-5	1000	186,15	40259,71	40392,00	0,33	107686	1800,08	20847 (11544/9018)	40392,00 (1)*	20,0	20,0	52,0
Euro-120-1	1000	315,70	40259,64	40436,00	0,44	63387	1800,06	12065 (6622/5162)	40436,00 (1)*	20,0	20,0	56,0
Euro-120-2	1000	325,90	40257,25	40502,00	0,60	65898	1800,24	12915 (6615/6029)	40502,00 (1)*	20,0	20,0	62,0
Euro-120-3	700	254,32	40258,17	40282,00	0,06	700	254,36	1 (1/0)	40282,00 (1)	20,0	20,0	42,0
Euro-120-4	1000	303,35	40259,81	40458,00	0,49	58425	1800,06	10949 (6235/4405)	40458,00 (1)*	20,0	20,0	58,0
Euro-120-5	1000	376,06	40259,03	50347,00	20,04	21850	1800,06	4253 (2126/1826)	42669,00 (276)*	21,0	21,0	78,0

^aUm asterisco indica a falha na obtenção da solução ótima

Tabela 5.11: Resultados do Algoritmo de *Branch-and-Bound* **BBR3** para Rede Pan-Européia

Instância	Nó Inicial					Árvore de Busca				ADMs	LPs	WLs
	Iter	Tempo (seg)	Limite Inferior	Limite Superior	Gap (%)	Iter	Tempo (seg)	Nós da Árvore Gerados(Expl./Podados)	Solução (Nó da Solução) ^a			
Euro-80-1	1000	125,80	40260,00	40348,00	0,22	15735	1800,40	2565 (1737/663)	40348,00 (1)*	20,0	20,0	48,0
Euro-80-2	1000	126,39	40260,00	40480,00	0,54	15280	1801,10	2859 (1429/1130)	40480,00 (1)*	20,0	20,0	60,0
Euro-80-3	1000	124,47	40260,00	40348,00	0,22	16141	1800,44	2473 (1972/431)	40348,00 (1)*	20,0	20,0	48,0
Euro-80-4	1000	123,22	40260,00	40348,00	0,22	15600	1800,07	2923 (1461/1153)	40348,00 (1)*	21,0	20,0	48,0
Euro-80-5	1000	126,23	40260,00	40326,00	0,16	15408	1800,79	2883 (1442/1138)	40326,00 (1)*	20,0	20,0	46,0
Euro-120-1	1000	188,54	40260,00	41382,00	2,71	5130	1800,94	829 (414/68)	41382,00 (1)*	20,0	20,0	142,0
Euro-120-2	1000	198,34	40260,00	41030,00	1,88	9070	1801,50	1617 (808/502)	41030,00 (1)*	20,0	20,0	110,0
Euro-120-3	1000	193,99	40260,00	40458,00	0,49	5530	1801,14	909 (454/88)	40458,00 (1)*	20,0	20,0	58,0
Euro-120-4	1000	191,28	40260,00	40524,00	0,65	8740	1800,66	1551 (775/451)	40524,00 (1)*	20,0	20,0	64,0
Euro-120-5	1000	189,70	40260,00	58795,00	31,52	6890	1801,00	1181 (590/213)	40744,00 (18)*	20,0	20,0	84,0

^aUm asterisco indica a falha na obtenção da solução ótima

Tabela 5.12: Tráfego do Anel com 4 Nós

Elemento de Origem	Elemento de Destino			
	A01	A02	A03	A04
A01	0	1	8	4
A02	12	0	3	9
A03	1	2	0	2
A04	4	1	7	0

e analisadas em [156] em que os autores fornecem a descrição completa de uma pequena instância de rede em anel SONET+WDM unidirecional cuja topologia é exibida na Figura 5.11. Em [156], utilizou-se a matriz de tráfego apresentada na Tabela 5.12 (cuja a unidade não foi fornecida) e considerou-se inicialmente uma taxa de *grooming* igual a três. As abordagens propostas e discutidas em [156] foram o uso de uma formulação de programação linear inteira (já descrita anteriormente na seção 2.4) em conjunto com um pacote comercial de otimização *CPLEX* (cuja versão não foi fornecida), uma heurística gulosa e um algoritmo baseado no uso de *simulated-annealing*. Os testes com o *CPLEX* foram realizados em uma estação de trabalho HP Visualize B1321 utilizando UNIX; enquanto que as outras duas abordagens foram testadas em um processador Pentium de 200 MHz utilizando Windows NT.

Segundo os autores em [156], após 6 horas de execução o *CPLEX* não foi capaz de encontrar uma solução ótima e o melhor resultado obtido utilizava 15 comprimentos de onda e 31 ADMs. O algoritmo de *simulated-annealing* foi capaz de produzir a mesma solução em apenas 2 segundos; enquanto que a heurística gulosa encontrou uma solução com 33 ADMs em um tempo desprezível.

Além disso, em [156] procurou-se comparar os resultados produzidos pelo algoritmo de *simulated-annealing* (*SA*) para o **TGP-SH** com aqueles produzidos por uma heurística gulosa (*HG*) para o **TGP-MH** proposta pelos mesmos autores em [33]. Para tanto foi utilizada a mesma matriz de tráfego (Tabela 5.12); todavia, a taxa de *grooming* variou entre os valores 3, 12 e 48. Os resultados obtidos são mostrados na Tabela 5.13, juntamente com os resultados produzidos pela aplicação da heurística lagrangeana proposta anteriormente (seção 5.5).

A primeira coluna representa a taxa de *grooming* (que para a formulação proposta neste texto é equivalente a fazer com que todos os produtos tenham o mesmo valor para m_p). A segunda e terceira coluna apresentam, respectivamente, o número de comprimentos de onda e de ADMs utilizados na solução gerada pelo algo-

Tabela 5.13: Comparação entre Abordagens

Taxa de <i>Grooming</i>	<i>SA</i> [156]		<i>HG</i> [33]		Heurística Lagrangeana				
	W	ADM	W	ADM	W	ADM	Gap (%)	Iter.	Tempo (seg.)
3	15	31	19	38	5	27	5,70	500	8,62
12	4	14	5	11	2	8	0,48	240	3,35
48	1	4	2	5	1	4	0,01	254	3,07

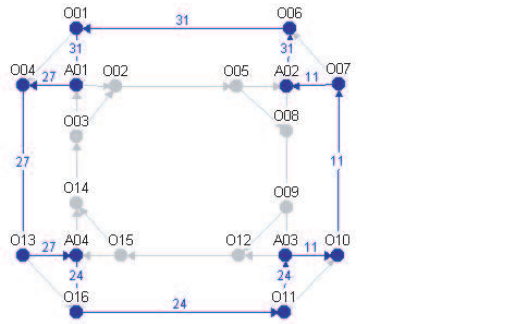


Figura 5.12: Solução Obtida para **TGP** com $m_p = 48$

ritmo de *simulated-annealing* (SA) para o **TGP-SH** [156]; enquanto que as duas colunas seguintes apresentam as mesmas informações para a solução obtida através da heurística gulosa (HG) para o **TGP-MH** [33]. As demais colunas sintetizam os resultados obtidos pela heurística lagrangeana. A sexta e sétima coluna exibem, respectivamente, o número de comprimentos de onda e de ADMs utilizados nas soluções obtidas pela heurística lagrangeana. Por fim, nas colunas seguintes, apresenta-se o *gap de dualidade*, o número de iterações e o tempo gasto na obtenção de tais soluções.

É fácil constatar a superioridade da abordagem lagrangeana em relação às demais. Além de produzir soluções utilizando um menor (ou igual) número de ADMs, as soluções geradas pela heurística lagrangeana utilizam um número menor (ou igual) de comprimentos de onda (mesmo sem conseguir provar a otimalidade da solução para o caso em que a taxa de *grooming* – ou m_p – era igual a 3). Apesar das diferenças intrínsecas a cada uma dessas abordagens (conforme discutido anteriormente), pode-se justificar o sucesso da abordagem lagrangeana da seguinte forma.

Em primeiro lugar, tanto a formulação proposta neste trabalho quanto a própria heurística lagrangeana parecem ser capazes de lidar melhor com o problema, na medida em que procuram explorar completa e eficientemente os recursos da rede (apesar do grande número de variáveis envolvidas). Isto pode ser observado pelas Figuras 5.12, 5.13 e 5.14, em que se apresentam as soluções obtidas pela heurística lagrangeana.

Cumpre destacar que a solução obtida para o problema (no caso em que a taxa

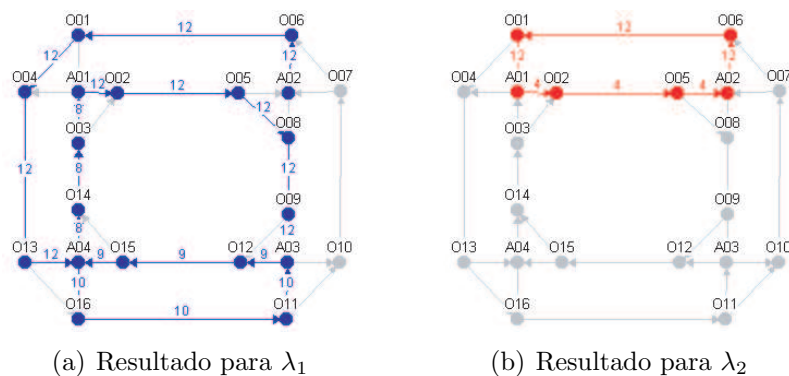


Figura 5.13: Solução Obtida para **TGP** com $m_p = 12$

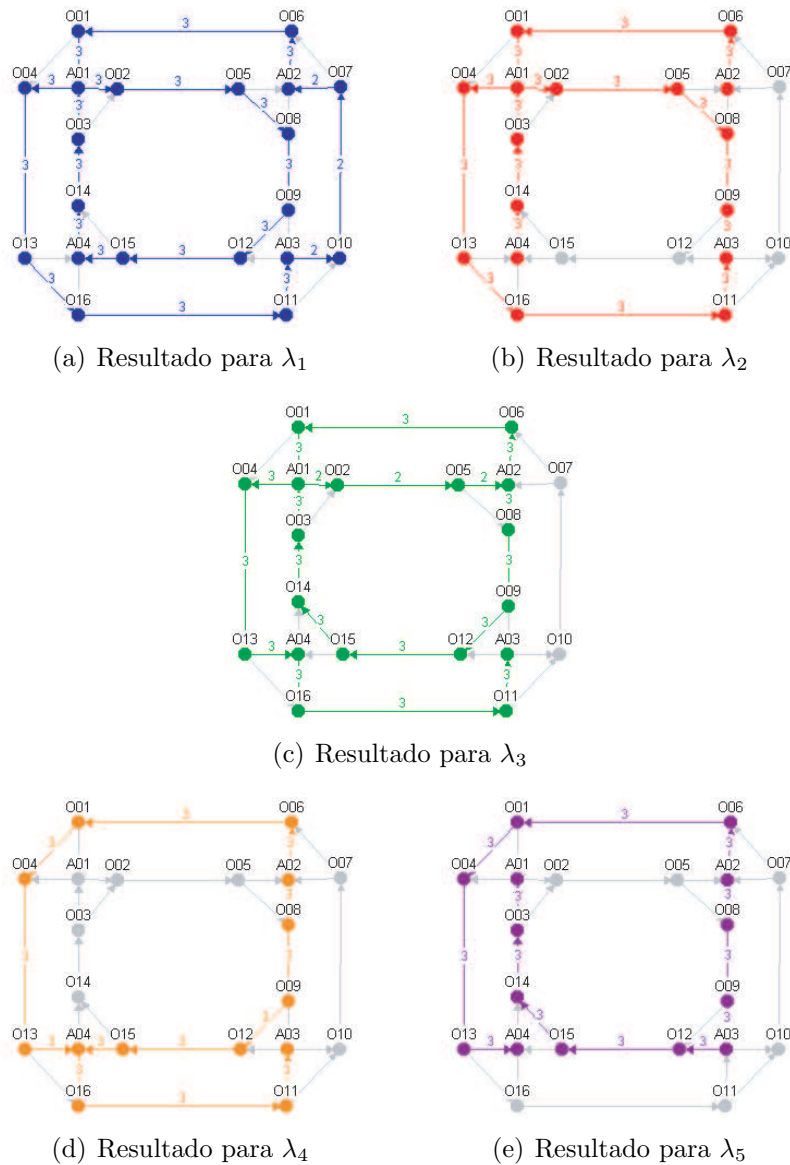


Figura 5.14: Solução Obtida para **TGP** com $m_p = 3$

de grooming era igual a 48) representa a adoção de uma topologia virtual equivalente a um simples anel unidirecional, ver a Figura 5.12, semelhante à solução gerada pelo algoritmo de *simulated-annealing*. Contudo, na medida em que a capacidade de transporte é reduzida (ou que o volume do tráfego aumenta), a heurística lagrangeana somente obteve resultados mais satisfatórios porque passou a utilizar o anel nos dois sentidos (horário e anti-horário) – ver Figuras 5.13 e 5.14. Infelizmente, resultados equivalentes para as outras duas abordagens não se encontram disponíveis, apesar de que, a princípio, elas não possuem nenhuma restrição ao fato de se considerar a rede como bidirecional (a não ser pelo grande número de variáveis – ver seção 2.4).

Outro aspecto importante que pode ajudar a justificar o bom desempenho da heurística lagrangeana está relacionado com a utilização da função de custo (3.7) que procura minimizar tanto o número de *lightpaths* (e, conseqüentemente, o número

de ADMs) como o comprimento dos mesmos (ver discussão no início do capítulo 3 e no final da subseção 3.1.2). Já nas outras duas abordagens, se prioriza apenas a minimização do número de ADMs (os próprios autores em [156] declaram o problema de se minimizar simultânea e eficientemente o número de comprimentos de onda juntamente com o número de ADMs como uma questão em aberto). De certa forma, parece que, ao se contabilizar o comprimento dos *lightpaths* na função objetivo, a heurística lagrangeana tende não só a selecionar rotas menores mas também a utilizá-las melhor (alocando mais tráfego àquelas já existentes), obtendo, neste caso, uma solução em que se explora melhor a capacidade dos *lightpaths*. Entretanto, tais afirmações não passam de simples conjecturas, sendo necessário se realizar uma investigação mais profunda para um melhor entendimento e embasamento dessas constatações.

Deve-se dizer que o intuito desta subseção é prover uma simples comparação entre a abordagem lagrangeana descrita neste trabalho e outras propostas encontradas na literatura para solucionar o **TGP**. A seguir, apresenta-se um conjunto maior de experimentos que foi realizado com o intuito de esclarecer melhor as questões levantadas nesta subseção.

5.7.2 Avaliação Qualitativa dos Resultados

A despeito das dificuldades levantadas na subseção anterior para realização de uma análise comparativa e em vista dos “bons” resultados obtidos na avaliação preliminar, torna-se necessária uma avaliação qualitativa mais profunda das abordagens propostas neste capítulo para que se tenha uma melhor compreensão das reais contribuições deste trabalho.

De modo a realizar tal avaliação, a heurística proposta em [166] foi implementada em C++ e testada utilizando-se a mesma plataforma que as abordagens descritas neste capítulo. A heurística descrita em [166] também utiliza um grafo auxiliar em camadas que é construído a partir da topologia da rede original. A Figura 5.15(b) exhibe um exemplo do grafo auxiliar para a topologia de rede apresentada na Figura 5.15(a).

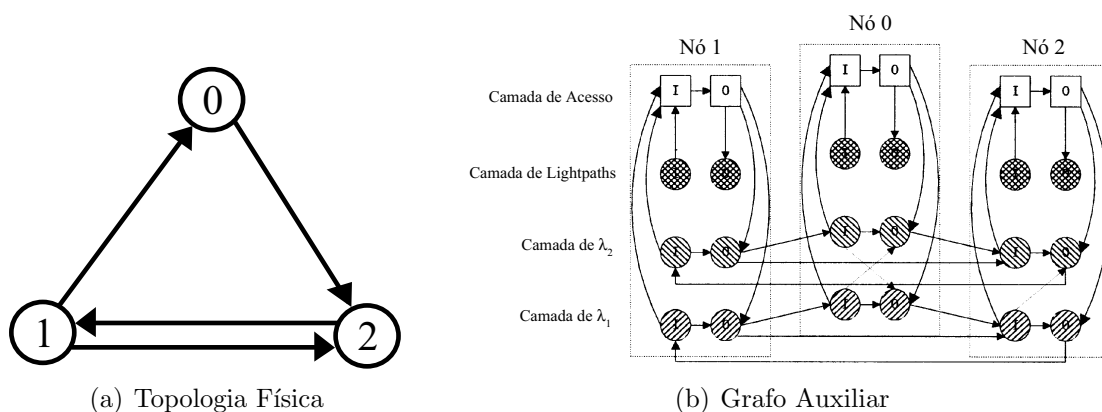


Figura 5.15: Exemplo de Grafo Auxiliar (adaptado de [166])

O grafo auxiliar utilizado em [166] possui $W + 2$ camadas. As camadas de 1 a W representam cada comprimento de onda separadamente. Já a camada $W + 1$ representa a camada de *lightpaths*, enquanto a camada $W + 2$ é denominada camada de acesso (de onde todo o tráfego é originado e terminado). Cada nó da rede possui duas portas em cada camada (uma de entrada e outra de saída). Além disso, o grafo auxiliar possui os seguintes arcos:

- **WBE** – *Wavelength Bypass Edges*: arcos que ligam cada porta de entrada de uma camada associada a um comprimento de onda a respectiva porta de saída da mesma camada para cada um dos nós;
- **GrmE** – *Grooming Edges*: arcos que ligam a porta de entrada da camada de acesso a respectiva porta de saída da mesma camada em cada um dos nós;
- **MuxE** – *Mux Edges*: arcos que ligam a porta de saída da camada de acesso à porta de saída da camada de *lightpaths* para cada um dos nós;
- **DmxE** – *Demux Edges*: arcos que ligam a porta de entrada da camada de *lightpaths* à porta de entrada da camada de acesso para cada um dos nós;
- **TxE** – *Transmitter Edges*: arcos que ligam a porta de saída da camada de acesso a cada uma das portas de saída das camadas associadas aos comprimentos de onda para cada um dos nós;
- **RxE** – *Receiver Edges*: arcos que ligam cada uma das portas de entrada das camadas associadas aos comprimentos de onda à porta de entrada da camada de acesso para cada um dos nós;
- **WLE** – *Wavelength-Link Edges*: arcos que ligam a porta de saída de uma camada associada a um comprimento de onda do nó i à porta de entrada da camada associada ao mesmo comprimento de onda do nó j , se existe o arco (conexão física) entre i e j ;
- **LPE** – *Lightpath Edges*: arcos que ligam a porta de saída de uma camada de *lightpaths* do nó i à porta de entrada da camada de *lightpaths* do nó j , se existe um *lightpath* entre i e j .

A heurística proposta em [166] atende uma requisição de tráfego por vez, utilizando um algoritmo de caminho mínimo no grafo auxiliar para:

- determinar uma rota na topologia virtual (usando arcos do tipo **LPE**) caso exista capacidade disponível nos *lightpaths* já estabelecidos; ou
- alocar novos *wavelength-links*, estabelecendo, assim, novos *lightpaths* que serão utilizados para transportar a requisição corrente.

Em [166], demonstra-se que diferentes políticas de *grooming* podem ser obtidas a partir da manipulação dos custos atribuídos aos arcos. Três políticas distintas são definidas e analisadas em [166]:

Tabela 5.14: Custos dos Arcos do Grafo Auxiliar para cada Política de *Grooming* [166]

Tipo do Arco	Políticas		
	MinTH	MinLP	MinWL
WLE	10	10	1000
GrmE	1000	20	0
TxE	20	200	20
RxE	20	200	20
LPE	1	1	1
MuxE	0	0	0
DmxE	0	0	0
WBE	0	0	0

- **MinTH**: política segundo a qual se procura transportar a requisição de tráfego corrente utilizando o número mínimo de “saltos” (*hops*) na topologia virtual;
- **MinLP**: política que procura estabelecer o menor número possível de novos *lightpaths* durante o atendimento da requisição de tráfego corrente;
- **MinWL**: política que procura alocar o menor número possível de novos *wavelength-links* de forma a atender a requisição de tráfego corrente.

A Tabela 5.14 apresenta os custos utilizados em [166] de modo a produzir os efeitos desejados em cada uma das políticas descritas anteriormente. Cumpre, ainda, destacar que em [166] são testados três esquemas distintos para seleção da requisição de tráfego a ser roteada em cada iteração. São eles:

- **LCF** – *Least Cost First*: este esquema seleciona a requisição cujo atendimento representa a solução mais efetiva em termos de custo. O custo de uma requisição é obtido dividindo-se o custo do caminho mínimo que seria utilizado para rotear o tráfego da requisição pelo total de tráfego a ser transportado (vale dizer que, após o roteamento de cada requisição, os custos das requisições ainda não atendidas devem ser recalculados, levando-se em consideração o novo estado da rede – isto é do grafo auxiliar);
- **MUF** – *Maximum Utilization First*: este esquema seleciona a requisição que possui a maior “utilização” (definida como o total de tráfego dividido pelo número de “saltos”, ou *hops*, entre a origem e o destino da requisição na topologia física);
- **MAF** – *Maximum Amount First*: este esquema seleciona a requisição que possui o maior valor de tráfego.

Por fim, os resultados em [166] mostram que o esquema LCF, apesar de ser computacionalmente mais caro, fornece melhores resultados que os demais.

De forma, a se comparar qualitativamente os resultados obtidos pelas abordagens propostas neste capítulo e a heurística proposta em [166], adotou-se uma estrutura

Tabela 5.15: Custos dos Arcos do Grafo Auxiliar da Política MinWL*

Tipo do Arco	Política MinWL*
WLE	11
GrmE	0
TxE	1001
RxE	1001
LPE	1
MuxE	0
DmxE	0
WBE	0

de custos para o grafo auxiliar semelhante àquela obtida pela utilização da equação (3.7). Essa “nova” política de *grooming* foi denominada **MinWL***, devido a semelhança de seus resultados com os da política **MinWL** (descrita anteriormente). A Tabela 5.15 apresenta os custos utilizados na implementação dessa política.

A heurística proposta em [166] foi aplicada as mesmas instâncias descritas anteriormente (150 problemas divididos em 3 conjuntos). Em tais testes utilizou-se, além das políticas de *grooming* descritas em [166], a política **MinWL***. A Tabela 5.16 apresenta os resultados obtidos para cada um dos subconjuntos das instâncias. A primeira coluna dessa tabela apresenta a identificação do subconjunto de instâncias. Em seguida, para cada uma das políticas, são apresentados os valores médios do custo total, do número de multiplexadores de *add-drop*, do número de *lightpaths* e do número de *wavelength-links*. É fácil de se constatar que os resultados apresentados pela “nova” política (**MinWL***) se assemelham aos obtidos pela política **MinWL** definida em [166]. Em verdade, para a maioria dos casos, ambas as políticas obtiveram, não só os melhores resultados para o número médio de *wavelength-links* alocados, mas também para o número médio de multiplexadores de *add-drop* e para o número médio de *lightpaths*.

Devido às diferenças na estrutura de custo utilizadas pelas políticas definidas em [166] e às propostas apresentadas neste capítulo, a comparação entre elas deve ser realizada de forma indireta (através da política **MinWL***). Em primeiro lugar, conforme observado anteriormente, os resultados obtidos por **MinWL*** superam aqueles obtidos pelas demais políticas definidas [166]. Tal constatação se baseia no fato de que o custo do projeto de rede virtual (conforme descrito anteriormente) está diretamente relacionado ao número de *lightpaths* e de ADMs (ou LTEs). Por outro lado, se tais resultados forem comparados aos obtidos por **BBR2** e **BBR3**, verifica-se uma clara superioridade da abordagem lagrangeana em relação a heurística proposta em [166].

Os gráficos da Figura 5.16 apresentam uma comparação entre os valores médios obtidos pela heurística proposta em [166] – utilizando a política **MinWL*** – e os resultados de **BBR2** e **BBR3** para as redes em anel. Observa-se que qualquer uma das abordagens lagrangeanas obtém valores 50% menores que os obtidos pela heurística proposta em [166] tanto para o número de multiplexadores de *add-drop* quanto para o número de *lightpaths*. Em verdade, as soluções obtidas pela heurística são

Tabela 5.16: Resultados da Heurística proposta em [166] (valores médios)

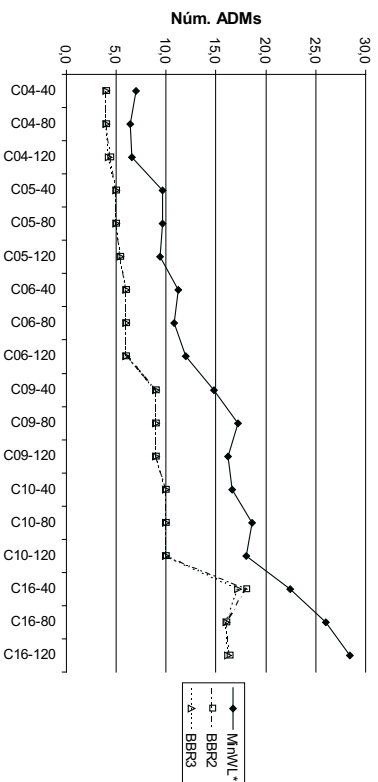
Instância	MinLP				MinWL				MinTH				MinWL*			
	Custo	ADMs	LPs	WLs	Custo	ADMs	LPs	WLs	Custo	ADMs	LPs	WLs	Custo	ADMs	LPs	WLs
C04-40	2624	7,0	6,4	6,4	6656	7,0	6,4	6,4	390	8,0	7,8	7,8	12883	7,0	6,4	6,4
C04-80	2378	6,4	5,8	5,8	6032	6,4	5,8	5,8	400	8,0	8,0	8,0	11675	6,4	5,8	5,8
C04-120	2542	6,6	6,2	6,2	6448	6,6	6,2	6,2	400	8,0	8,0	8,0	12481	6,6	6,2	6,2
C05-40	3198	9,6	7,8	7,8	8112	9,6	7,8	7,8	500	10,0	10,0	10,0	15701	9,6	7,8	7,8
C05-80	3362	9,6	8,2	8,2	8528	9,6	8,2	8,2	500	10,0	10,0	10,0	16507	9,6	8,2	8,2
C05-120	3362	9,4	8,2	8,2	8528	9,4	8,2	8,2	500	10,0	10,0	10,0	16507	9,4	8,2	8,2
C06-40	4100	11,2	10,0	10,0	10400	11,2	10,0	10,0	590	12,0	11,8	11,8	20130	11,2	10,0	10,0
C06-80	4018	10,8	9,8	9,8	10192	10,8	9,8	9,8	600	12,0	12,0	12,0	19727	10,8	9,8	9,8
C06-120	4264	12,0	10,4	10,4	10816	12,0	10,4	10,4	600	12,0	12,0	12,0	20935	12,0	10,4	10,4
C09-40	5606	14,4	13,6	16,6	16544	14,8	13,6	16,0	788	15,2	15,2	18,0	27403	14,8	13,6	16,0
C09-80	6480	17,2	15,8	16,0	16632	17,2	15,8	16,0	876	17,4	17,4	18,0	31808	17,2	15,8	16,0
C09-120	5822	16,2	14,2	14,2	14768	16,2	14,2	14,2	876	17,4	17,4	18,0	28585	16,2	14,2	14,2
C10-40	6332	16,6	15,4	17,2	17816	16,6	15,4	17,2	900	18,0	17,6	19,6	31020	16,6	15,4	17,2
C10-80	7058	18,6	17,2	17,8	18488	18,6	17,2	17,8	960	19,0	19,0	20,0	34630	18,6	17,2	17,8
C10-120	6728	18,4	16,4	16,8	18256	18,0	16,4	17,6	976	19,4	19,4	20,0	33026	18,0	16,4	17,6
C16-40	8520	22,0	20,6	28	28032	22,4	20,8	27,2	1218	23,0	22,6	31,4	41941	22,4	20,8	27,2
C16-80	9900	25,6	24,0	30,0	29776	26,0	24,4	28,8	1376	26,4	26,4	32,0	49177	26,0	24,4	29,8
C16-120	10848	28,0	26,4	28,8	29456	28,4	26,4	28,4	1464	28,6	28,6	32,0	53165	28,4	26,4	28,4
M06-40	4434	13,2	10,8	11,4	11232	13,2	10,8	10,8	712	15,4	14,2	14,4	21740	13,2	10,8	10,8
M06-80	4182	11,2	10,2	10,2	10608	11,2	10,2	10,2	800	16,0	16,0	16,0	20533	11,2	10,2	10,2
M06-120	4264	12,0	10,4	10,4	10816	12,0	10,4	10,4	800	16,0	16,0	16,0	20935	12,0	10,4	10,4
M10-40	6664	18,8	16,2	18,4	16408	17,4	15,2	15,8	986	20,6	19,2	21,8	30604	17,2	15,2	15,8
M10-80	6992	19,4	17,0	19,2	17672	19,0	16,8	17,0	1096	22,4	21,6	23,2	33821	19,2	16,8	17,0
M10-120	7304	20,4	17,8	18,4	18912	20,4	17,8	18,2	1112	22,8	22,0	23,2	35836	20,4	17,8	18,2
M15-80	10732	29,4	26,0	33,2	27192	28,2	24,8	26,2	1562	30,8	29,0	40,2	49938	28,2	24,8	26,2
M15-120	10872	30,0	26,4	31,2	27440	29,0	26,0	26,4	1824	35,6	35,2	41,6	52342	29,0	26,0	26,4
N20-80	12310	32,8	29,8	39,0	36832	35,4	30,8	35,6	1798	34,2	33,2	47,0	60051	34,6	29,8	35,6
N20-120	13848	37,0	33,6	40,8	37520	36,8	33,0	36,2	1988	38,2	37,8	47,6	66464	36,8	33,0	36,2
Euro-80	14866	40,6	35,6	62,6	38720	38,2	33,0	37,4	2438	45,2	42,6	73,4	66086	37,6	32,8	38,2
Euro-120	14918	41,8	36,0	51,8	38592	40,4	34,8	37,2	2822	54,6	51,8	75,0	70079	40,4	34,8	37,2

de qualidade igual àquela obtida pela heurística usada para inicialização do limite superior do processo de busca. O bom desempenho das abordagens lagrangeanas está relacionado ao fato de que durante o processo de exploração (tanto do nó inicial quanto dos demais) a atualização dos multiplicadores de Lagrange facilita a obtenção de novas soluções viáveis de melhor qualidade, que se utilizam de uma quantidade menor de recursos (multiplexadores de *add-drop* e *lightpaths*). A heurística proposta em [166], além de examinar uma única vez cada requisição, também se mostra incapaz de modificar as alocações já realizadas anteriormente; dificultando, assim, a obtenção de melhores soluções.

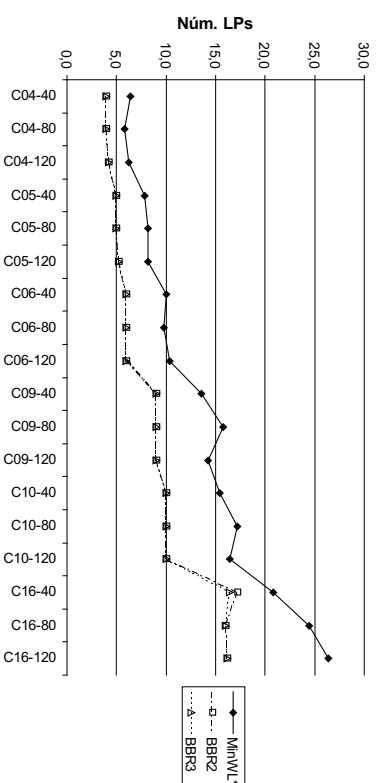
Além disso, os resultados obtidos pelas abordagens lagrangeanas podem ser melhor interpretados examinando-se o número médio de *wavelength-links* alocados nas soluções obtidas por **BBR2** e **BBR3**. Pelo gráfico da Figura 5.16(c), constata-se que o número médio de *wavelength-links* alocados pelas duas abordagens lagrangeanas é praticamente o mesmo (exceto para dois subconjuntos das instâncias com 16 nós – aquelas com 40 e com 120 produtos). Além disso, os valores médios obtidos pelas duas abordagens lagrangeanas superam àqueles obtidos pela heurística proposta em [166]. Em verdade, ao se examinar o gráfico da Figura 5.16(d), verifica-se que o número médio de “saltos” (*hops*) na topologia virtual obtida pelas abordagens lagrangeanas chega a ser de duas a três vezes maior que aquele obtido pela heurística. Isso conduz a conclusão de que o motivo das abordagens lagrangeanas conseguirem explorar melhor os recursos da rede (obtendo soluções com um número menor de ADMs e *lightpaths*) está ligado ao fato delas utilizarem mais “saltos” (*hops*) na topologia virtual para rotear as requisições de tráfego. Por outro lado, a heurística proposta em [166] parece optar, preferencialmente, pelo roteamento com um “único salto” (*single hop*); justificando, assim, a sua maior necessidade de ADMs e *lightpaths* para transportar o mesmo conjunto de requisições.

Já os gráficos da Figura 5.17 apresentam uma comparação entre os valores médios obtidos pela heurística proposta em [166] – utilizando a política **MinWL*** – e os resultados de **BBR2** e **BBR3** para as redes irregulares. Novamente, os resultados obtidos por **BBR2** e **BBR3** para os valores médios de ADMs e de *lightpaths* superam (isto é, são menores) que os obtidos pela heurística.

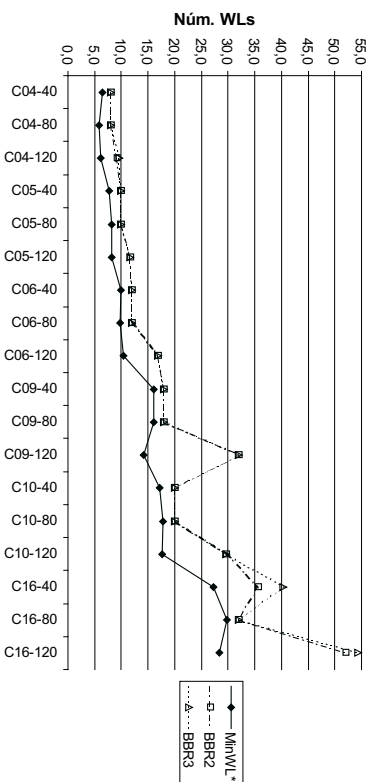
Mais ainda, analisando-se os gráficos das Figuras 5.17(c) e 5.17(d), observa-se o mesmo padrão encontrado para as redes em anel. Contudo, enquanto que para **BBR2** o número médio de “saltos” (*hops*) na topologia virtual continua a ser de duas a três vezes maior que aquele obtido pela heurística, para **BBR3** esse número chega a ser de quatro a cinco vezes maior. Este comportamento pode ser atribuído, pelo menos em parte, ao fato (já analisado anteriormente) de que a relaxação #3 piora seu desempenho (isto é, a qualidade das soluções obtidas) quando aplicada a grandes redes irregulares. Como visto anteriormente, isto se deve a forma como é resolvido o subproblema envolvendo as variáveis de fluxo (que está intimamente ligado ao roteamento do tráfego). Na relaxação #3, esse subproblema é aproximado de maneira “simplista” através de um procedimento de resolução por inspeção, enquanto que na relaxação #2 a aproximação do subproblema de fluxo é feita através do cálculo de vários caminhos mínimos. Essa estratégia, apesar de ser computacionalmente mais cara, apresenta melhores resultados tanto em relação ao número médio de *wavelength-links* alocados quanto em relação ao número de “saltos” (*hops*).



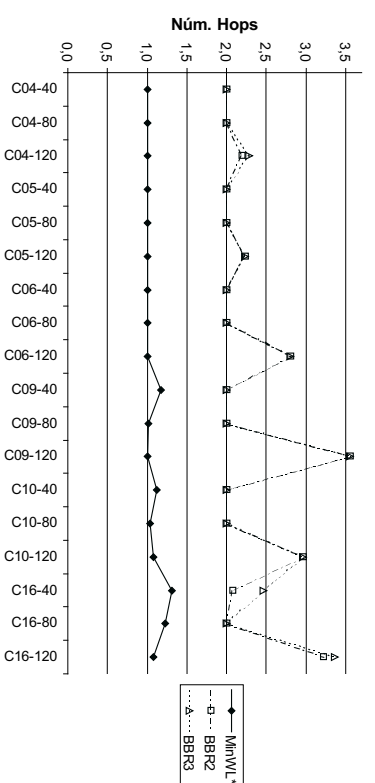
(a) Comparação do Número de ADMs



(b) Comparação do Número de LPs

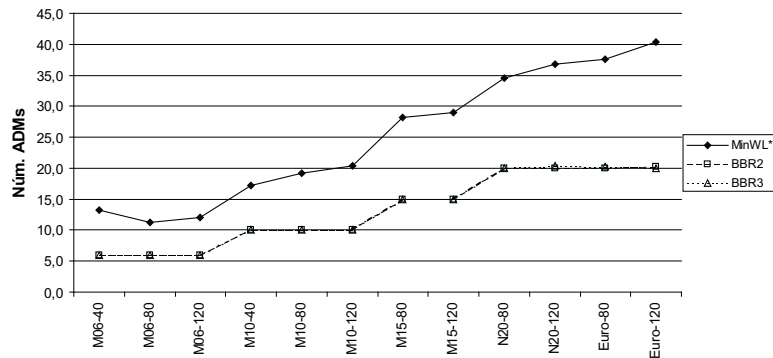


(c) Comparação do Número de WLS

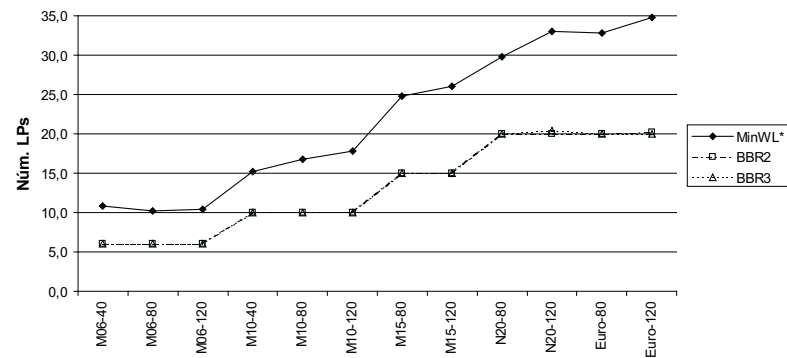


(d) Comparação do Número de Hops

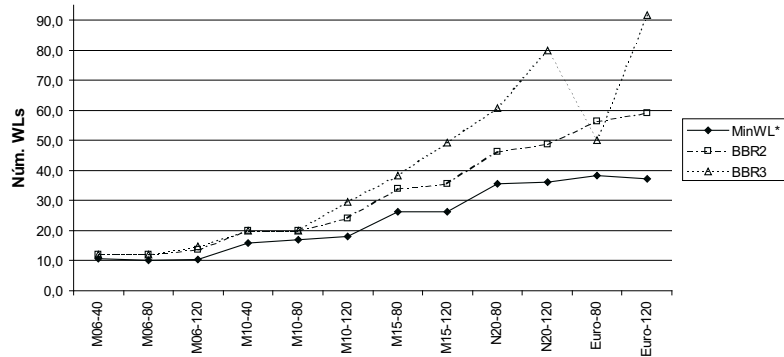
Figura 5.16: Comparação dos Resultados Médios dos Algoritmos de *Branch-and-Bound* para Redes em Anel



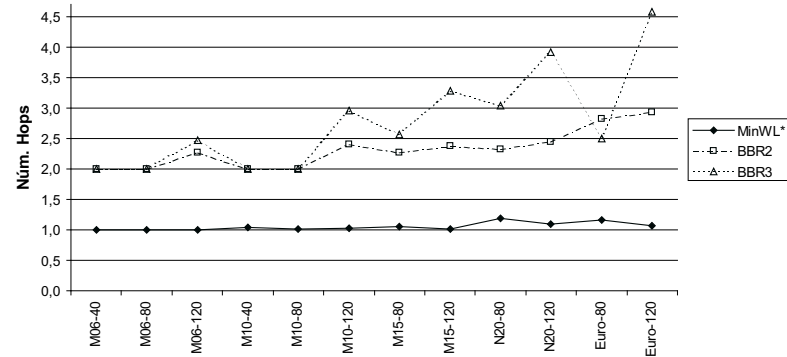
(a) Comparação do Número de ADMs



(b) Comparação do Número de LPs



(c) Comparação do Número de WLS



(d) Comparação do Número de Hops

Figura 5.17: Comparação dos Resultados Médios dos Algoritmos de *Branch-and-Bound* para Redes Irregulares

Finalmente, os gráficos das Figuras 5.18 e 5.19 apresentam uma comparação dos valores médios obtidos pela heurística lagrangeana com aqueles obtidos pela heurística proposta em [166] – utilizando a política **MinWL***, bem como com os resultados dos algoritmos de *branch-and-bound* **BBR2** e **BBR3** para as redes em anel e para as redes irregulares, respectivamente. Nesses gráficos, os resultados obtidos pela heurística lagrangeana estão indicados na legenda pelas siglas **HL-R2** e **HL-R3** de acordo com a relaxação utilizada (relaxação #2 ou #3).

Constata-se que, conforme já foi mencionado anteriormente ao final da subseção 5.6.2, as soluções obtidas pela heurística lagrangeana apresentam resultados praticamente idênticos àqueles obtidos pelos algoritmos de *branch-and-bound* utilizando as mesmas relaxações, exceto para a relaxação #3 quando aplicada a grandes instâncias das redes irregulares (caso em que essa relaxação #3 apresenta dificuldades já discutidas e analisadas anteriormente). Nesse último caso, apesar das soluções obtidas pela heurística lagrangeana também utilizarem praticamente o mesmo número de multiplexadores de *add-drop* e *lightpaths*, o procedimento de busca parece ser realmente necessário para se obter soluções com um número menor de “saltos” (*hops*) – ver gráfico da Figura 5.19(d).

5.8 Considerações Finais

Neste capítulo apresentou-se uma abordagem através da relaxação lagrangeana para resolução do **TGP**. Além de se formalizar o uso da relaxação lagrangeana através de uma descrição detalhada de quatro limites inferiores obtidos a partir de sua aplicação, procurou-se descrever os detalhes pertinentes à implementação do algoritmo de *branch-and-bound* baseado no uso da relaxação lagrangeana para tais problemas. Em particular, essa descrição abordou uma heurística para o cálculo de limites superiores e o detalhamento do método de subgradientes e do procedimento de busca na árvore de soluções.

Apresentou-se também uma heurística lagrangeana para o **TGP** utilizando as relaxações propostas, bem como os resultados de sua aplicação na resolução de mais de uma centena de instâncias do mesmo. Os resultados obtidos demonstram a superioridade das relaxações #2 e #3 em relação as demais. Além disso, a relaxação #2 parece ser mais adequada a resolução de grandes instâncias de redes irregulares.

Os resultados obtidos por duas implementações do algoritmo de *branch-and-bound* – **BBR2** e **BBR3**, para mais de uma centena de instâncias também foram apresentados e analisados, resultando na confirmação do desempenho superior da relaxação #2.

De modo a se avaliar qualitativamente as soluções obtidas pelas implementações dos algoritmos de *branch-and-bound* – **BBR2** e **BBR3**, implementou-se a heurística proposta em [166]. Os resultados obtidos apontam para a superioridade da abordagem lagrangeana, em especial da relaxação #2, apesar de seu maior custo computacional.

Contudo, vale dizer que **BBR2** foi capaz de obter e provar a otimalidade de 88% das instâncias, ao passo que para **BBR3** esse número se reduziu a 68%. Isto, aliado ao fato de que a maior parte das soluções foi obtida na exploração do nó inicial, leva a crer que “boas” soluções poderiam ser obtidas a partir da restrição

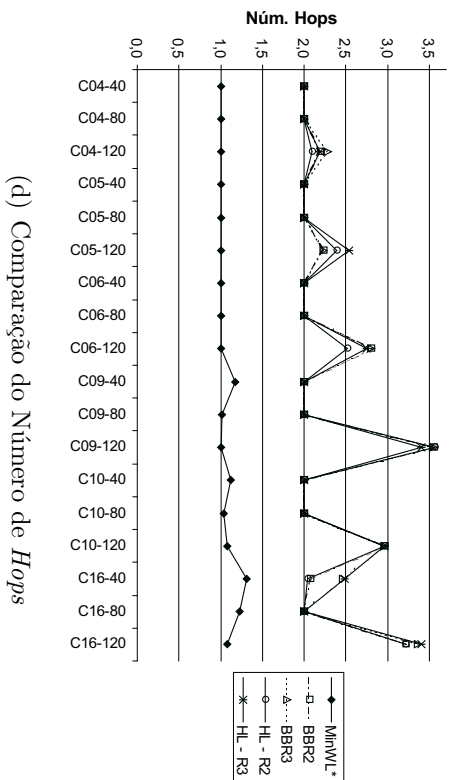
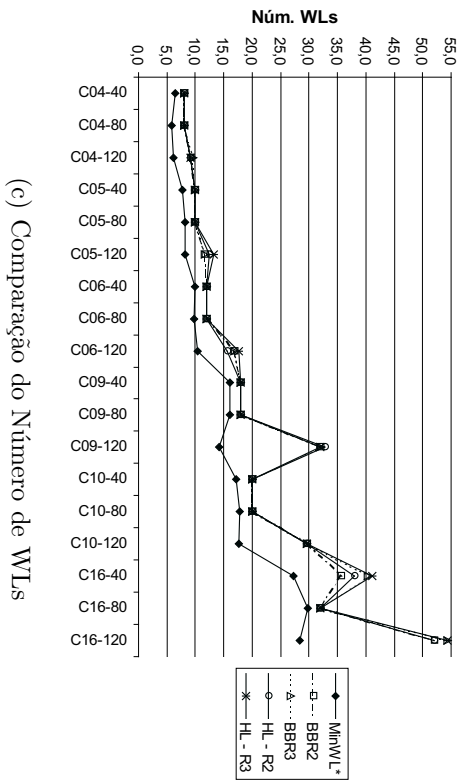
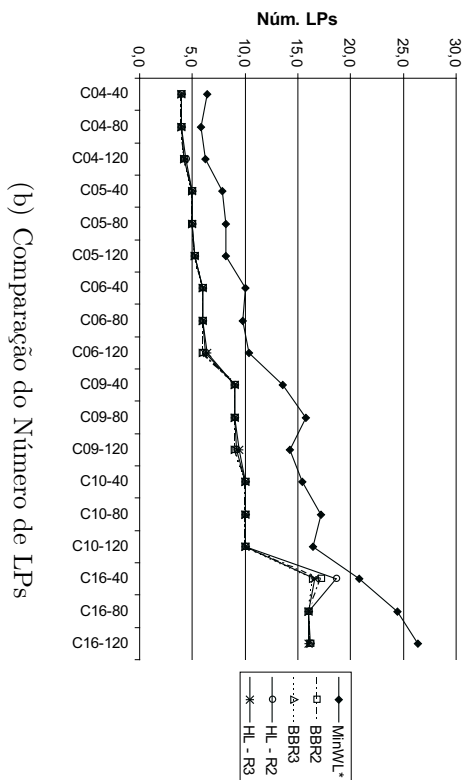
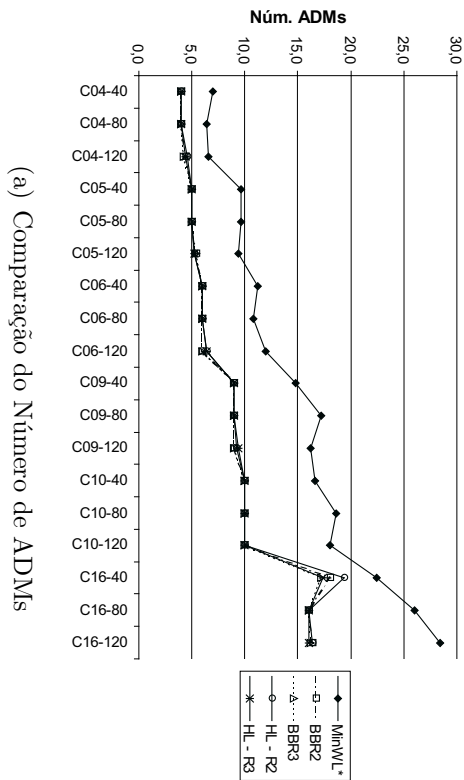
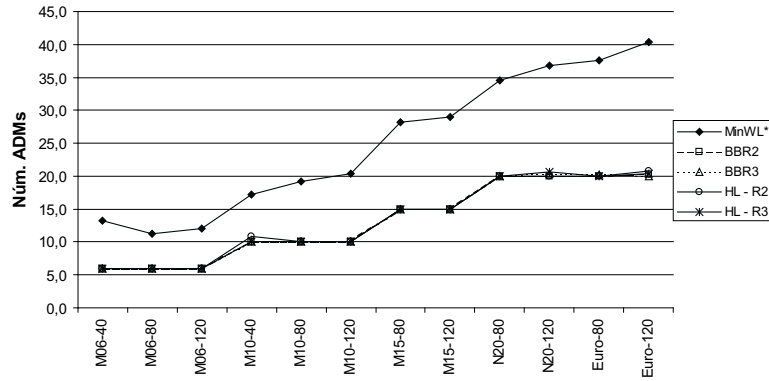
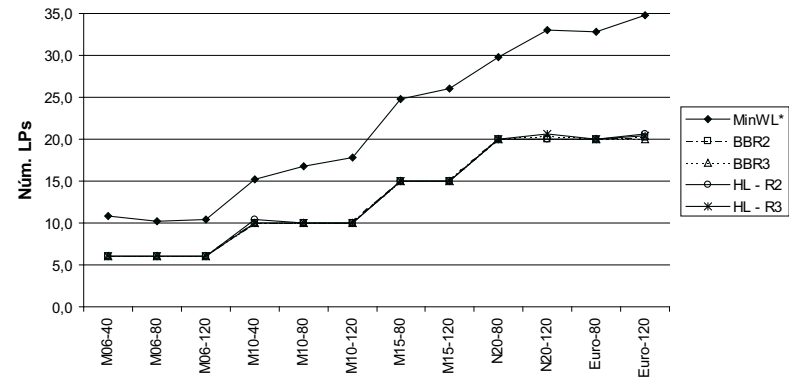


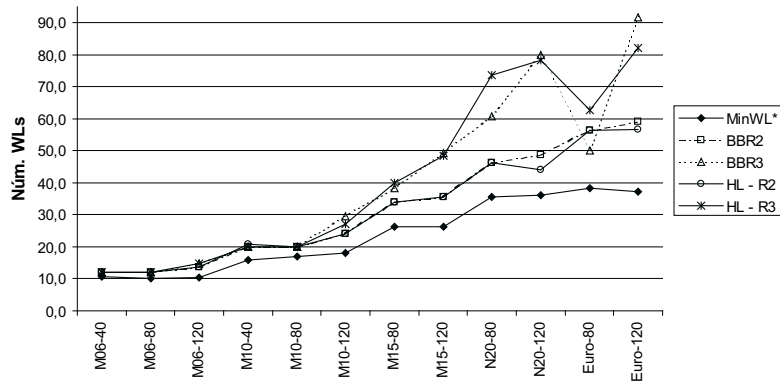
Figura 5.18: Comparação dos Resultados Médios da Heurística Lagrangeana para Redes em Anel



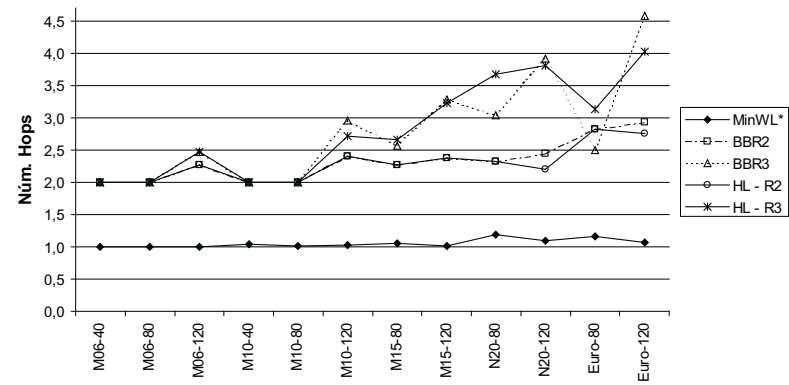
(a) Comparação do Número de ADMs



(b) Comparação do Número de LPs



(c) Comparação do Número de WLS



(d) Comparação do Número de Hops

Figura 5.19: Comparação dos Resultados Médios da Heurística Lagrangeana para Redes Irregulares

do processo de busca apenas à exploração do nó inicial (o que equivale, à aplicação apenas da heurística lagrangeana) ou, ainda, da aplicação de uma forma limitada do mesmo. Na verdade, a comparação entre os resultados obtidos pela heurística lagrangeana utilizando as relaxações #2 e #3 e os resultados dos algoritmos de *branch-and-bound* **BBR2** e **BBR3** demonstra a eficácia da heurística lagrangeana na obtenção de “boas” soluções para o problema (utilizando apenas uma pequena fração do tempo necessário aos algoritmos de *branch-and-bound*).

No próximo capítulo, investiga-se o poliedro associado ao conjunto de soluções do **TGP** e os resultados obtidos a partir de tal estudo são aplicados na obtenção de soluções para o **TGP** através de um algoritmo de *branch-and-cut*.

Capítulo 6

Abordagem Poliédrica do TGP

NESTE capítulo são apresentados os resultados obtidos pela adoção de uma abordagem poliédrica para resolução do **TGP**. Na próxima seção, apresenta-se uma investigação do poliedro associado ao conjunto de soluções do **TGP** juntamente com uma análise dos poliedros associados a subestruturas da formulação do problema TGP_C apresentada na subseção 3.1.4. Já a seção 6.2 apresenta os resultados de uma avaliação preliminar de alguns cortes obtidos a partir desse estudo (apresentados, também, em [121]). As seções 6.3, 6.4 e 6.5 descrevem detalhadamente as implementações do método de *planos-de-corte*, do algoritmo de *branch-and-cut* e do algoritmo de *cut-and-branch*, respectivamente. Os resultados obtidos nos experimentos computacionais são apresentados na seção 6.6. Por fim, a seção 6.7 apresenta algumas considerações finais sobre essa abordagem.

6.1 Sobre o Poliedro de Soluções – \mathcal{P}_S

Seja \mathcal{S} o conjunto de soluções de um problema, então \mathcal{P}_S representa o poliedro formado pela envoltória convexa das soluções do problema. Dessa forma, para a formulação dada, $\mathcal{S} = \{(w, f) \in \{0, 1\}^{|A_\ell|} \times \mathbb{Z}^{|A_\ell| \times |P|} \mid (w, f) \text{ satisfaz a (3.10b)–(3.10g)}\}$, e o poliedro associado a \mathcal{S} , \mathcal{P}_S , é dado por:

$$\mathcal{P}_S \equiv \text{conv} \{(w, f) \in \{0, 1\}^{|A_\ell|} \times \mathbb{Z}^{|A_\ell| \times |P|} \mid (w, f) \text{ satisfaz a (3.10b)–(3.10g)}\}.$$

Considere, ainda, o poliedro \mathcal{LP}_S associado a relaxação linear das variáveis de fluxo, ou melhor,

$$\mathcal{LP}_S \equiv \text{conv} \{(w, f) \in \{0, 1\}^{|A_\ell|} \times \mathbb{R}^{|A_\ell| \times |P|} \mid (w, f) \text{ satisfaz a (3.10b)–(3.10f)}\}.$$

Como \mathcal{LP}_S foi obtido a partir de uma relaxação do modelo original, isto é, $\mathcal{P}_S \subseteq \mathcal{LP}_S$, qualquer inequação válida para o problema relaxado também será válida para o problema original.

Antes de se analisar o poliedro associado ao modelo proposto, deve-se introduzir duas formulações baseadas na relaxação linear das variáveis de fluxo do modelo original. Uma vez relaxadas as restrições (3.10g) responsáveis pela integralidade dos fluxos, na primeira formulação irá se substituir tais variáveis (f_ℓ^p) por um conjunto

de variáveis $x_\ell^p \in \mathbb{R}$, representando a ocupação da capacidade do arco ℓ pelo fluxo f_ℓ^p , isto é, $x_\ell^p = \delta_p f_\ell^p, \forall t_p \in P, \forall \ell \in A_\ell$ e δ_p dado por (3.2).

Dessa forma, obtém uma formulação modificada, $TGPM$, dada por:

$$(TGPM) \quad \min \sum_{\ell \in A_\ell} c_\ell w_\ell \quad (6.1a)$$

sujeito a:

$$\sum_{\ell \in A_\ell^+(i)} x_\ell^p - \sum_{\ell \in A_\ell^-(i)} x_\ell^p = \bar{b}_i^p, \quad \forall t_p \in P, \forall i \in N_\ell \quad (6.1b)$$

$$\sum_{\ell \in A_\ell^+(i)} w_\ell - \sum_{\ell \in A_\ell^-(i)} w_\ell = 0, \quad \forall i \in N_\ell^o \quad (6.1c)$$

$$\sum_{t_p \in P} x_\ell^p \leq \Delta w_\ell, \quad \forall \ell \in A_\ell \quad (6.1d)$$

$$x_\ell^p \geq 0, \quad \forall t_p \in P, \forall \ell \in A_\ell \quad (6.1e)$$

$$w_\ell \in \{0, 1\}, \quad \forall \ell \in A_\ell \quad (6.1f)$$

em que $\bar{b}_i^p = \delta_p b_i^p$, ou ainda:

$$\bar{b}_i^p = \begin{cases} \delta_p f_p & , \text{ se } i = s_p \\ -\delta_p f_p & , \text{ se } i = d_p \\ 0 & , \text{ se } i \neq s_p \neq d_p \end{cases} \quad \forall t_p \in P, \forall i \in N_\ell. \quad (6.2)$$

Essa formulação é totalmente equivalente à formulação relaxada, isto é, à formulação original desconsiderando as restrições de integralidade dos fluxos. O que equivale dizer que para toda solução (w, x) de $TGPM$ pode-se calcular uma solução $(w, f) \in \mathcal{LP}_S$, bastando para tanto fazer $f_\ell^p = x_\ell^p / \delta_p, \forall t_p \in P, \forall \ell \in A_\ell$. Seja \mathcal{M} o conjunto de soluções da formulação modificada, isto é,

$$\mathcal{M} = \{(w, x) \in \{0, 1\}^{|A_\ell|} \times \mathbb{R}^{|A_\ell| \times |P|} \mid (w, x) \text{ satisfaz a (6.1b)–(6.1f)}\}$$

e o poliedro associado a \mathcal{M} , \mathcal{P}_M , é dado por:

$$\mathcal{P}_M \equiv \text{conv} \{(w, x) \in \{0, 1\}^{|A_\ell|} \times \mathbb{R}^{|A_\ell| \times |P|} \mid (w, x) \text{ satisfaz a (6.1b)–(6.1f)}\}.$$

Pode-se, então, estabelecer o seguinte lema cuja demonstração é óbvia.

Lema 6.1. *Para os poliedros \mathcal{LP}_S e \mathcal{P}_M pode-se afirmar que:*

(a) $(w, f) \in \mathcal{LP}_S \iff (w, x) \in \mathcal{P}_M$, em que $x_\ell^p = \delta_p f_\ell^p, \forall t_p \in P, \forall \ell \in A_\ell$,

(b) $\min \{c^T w : (w, f) \in \mathcal{LP}_S\} = \min \{c^T w : (w, x) \in \mathcal{P}_M\}$,

(c) $a^T w + b^T f \geq \alpha$ é válida para $\mathcal{LP}_S \iff \sum_{\ell \in A_\ell} a_\ell w_\ell + \sum_{t_p \in P} \frac{1}{\delta_p} \sum_{\ell \in A_\ell} b_\ell^p x_\ell^p \geq \alpha$ é válida para \mathcal{P}_M .

Em alguns casos, pode ser vantajoso introduzir variáveis extras em uma formulação. Isto pode conduzir a uma eventual redução do número de restrições e/ou resultar em uma simplificação substancial na forma das inequações/equações utilizadas originalmente. Neste caso específico uma extensão de $TGPM$ permitirá a

explicitação das relações entre o poliedro \mathcal{P}_M e outros poliedros associados a problemas já estudados e discutidos na literatura. Em decorrência do Lema 6.1, é possível estender os resultados obtidos para \mathcal{P}_M para o poliedro \mathcal{LP}_S e, em seguida, para o poliedro \mathcal{P}_S .

Sendo assim, a formulação $TGPM$ pode ser estendida através da introdução de dois conjuntos de variáveis: as variáveis $y_\ell^p \in \mathbb{R}, 0 \leq y_\ell^p \leq 1$, indicando a fração da demanda \bar{f}_p do produto t_p transportada através do arco ℓ , de modo que $x_\ell^p = \bar{f}_p y_\ell^p, \forall t_p \in P, \forall \ell \in A_\ell$, em que $\bar{f}_p = \delta_p f_p, \forall t_p \in P$; e as variáveis $z_\ell \in \mathbb{Z}$ representando a capacidade total instalada no arco ℓ , isto é, $z_\ell = \Delta w_\ell, \forall \ell \in A_\ell$ e Δ dado por (3.1).

Por fim, uma extensão $TGPE$ da formulação modificada $TGPM$ é por:

$$(TGPE) \quad \min \sum_{\ell \in A_\ell} c_\ell w_\ell \quad (6.3a)$$

sujeito a:

$$\sum_{\ell \in A_\ell^+(i)} x_\ell^p - \sum_{\ell \in A_\ell^-(i)} x_\ell^p = \bar{b}_i^p, \quad \forall t_p \in P, \forall i \in N_\ell \quad (6.3b)$$

$$x_\ell^p = \bar{f}_p y_\ell^p, \quad \forall t_p \in P, \forall \ell \in A_\ell \quad (6.3c)$$

$$\sum_{t_p \in P} \bar{f}_p y_\ell^p \leq z_\ell, \quad \forall \ell \in A_\ell \quad (6.3d)$$

$$z_\ell = \Delta w_\ell, \quad \forall \ell \in A_\ell \quad (6.3e)$$

$$\sum_{\ell \in A_\ell^+(i)} w_\ell - \sum_{\ell \in A_\ell^-(i)} w_\ell = 0, \quad \forall i \in N_\ell^o \quad (6.3f)$$

$$x_\ell^p \geq 0, \quad \forall t_p \in P, \forall \ell \in A_\ell \quad (6.3g)$$

$$y_\ell^p \geq 0, \quad \forall t_p \in P, \forall \ell \in A_\ell \quad (6.3h)$$

$$y_\ell^p \leq 1, \quad \forall t_p \in P, \forall \ell \in A_\ell \quad (6.3i)$$

$$z_\ell \in \mathbb{Z}, \quad \forall \ell \in A_\ell \quad (6.3j)$$

$$w_\ell \in \{0, 1\}, \quad \forall \ell \in A_\ell \quad (6.3k)$$

Definição 6.2. Tomando-se $TGPE$ como base, pode-se definir os seguintes conjuntos de soluções:

$$\mathcal{X} \equiv \{x \in \mathbb{R}^{|A_\ell| \times |P|} \mid x \text{ satisfaz a (6.3b) e (6.3g)}\},$$

$$\mathcal{Y} \equiv \{y \in \mathbb{R}^{|A_\ell| \times |P|} \mid \bar{f}_p y \in \mathcal{X} \text{ e } y \text{ satisfaz a (6.3h) e (6.3i)}\},$$

$$\mathcal{Z} \equiv \{z \in \mathbb{Z}^{|A_\ell|} \mid \exists y \in \mathcal{Y} \text{ e } (y, z) \text{ satisfaz a (6.3d)}\}, \text{ e}$$

$$\mathcal{W} \equiv \{w \in \{0, 1\}^{|A_\ell|} \mid \Delta w \in \mathcal{Z} \text{ e satisfaz a (6.3f)}\}.$$

os quais estão associados aos seguintes poliedros, respectivamente:

$$\mathcal{P}_X \equiv \text{conv} \{x \in \mathbb{R}^{|A_\ell| \times |P|} \mid x \text{ satisfaz a (6.3b) e (6.3g)}\},$$

$$\mathcal{P}_Y \equiv \text{conv} \{y \in \mathbb{R}^{|A_\ell| \times |P|} \mid \bar{f}_p y \in \mathcal{X} \text{ e } y \text{ satisfaz a (6.3h) e (6.3i)}\},$$

$$\mathcal{P}_Z \equiv \text{conv} \{z \in \mathbb{Z}^{|A_\ell|} \mid \exists y \in \mathcal{Y} \text{ e } (y, z) \text{ satisfaz a (6.3d)}\}, \text{ e}$$

$$\mathcal{P}_W \equiv \text{conv} \{w \in \{0, 1\}^{|A_\ell|} \mid \Delta w \in \mathcal{Z} \text{ e satisfaz a (6.3f)}\}.$$

Esses poliedros se relacionam uns com os outros e, em particular, com \mathcal{P}_M . Além disso, existe uma forma canônica (transformação linear) para se mapear \mathcal{P}_W em \mathcal{P}_Z , bem como \mathcal{P}_Y em \mathcal{P}_X , e $\mathcal{P}_Z \subseteq \mathcal{P}_Y$. Isso resulta no seguinte lema.

Lema 6.3. *Para os poliedros \mathcal{P}_X , \mathcal{P}_Y , \mathcal{P}_Z e \mathcal{P}_W pode-se afirmar que:*

- (a) $\min \{c^T w : (w, x) \in \mathcal{P}_M\} = \min \{c^T w : w \in \mathcal{P}_W\}$,
- (b) $a^T w \geq \alpha$ é válida para $\mathcal{P}_M \iff a^T w \geq \alpha$ é válida para \mathcal{P}_W ,
- (c) Se $a^T z \geq \alpha$ é válida para $\mathcal{P}_Z \implies \sum_{\ell \in A_\ell} a_\ell w_\ell \geq \alpha/\Delta$ é válida para \mathcal{P}_W ,
- (d) Se $a^T y \geq \alpha$ é válida para \mathcal{P}_Y e $a_\ell^* > 0, \forall \ell \in A_\ell$, em que $a_\ell^* = \max_{t_p \in P} \{a_\ell^p\} \implies \sum_{\ell \in A_\ell} a_\ell^* z_\ell \geq \alpha f^*$ é válida para \mathcal{P}_Z , em que $f^* = \min_{t_p \in P} \{\bar{f}_p\}$,
- (e) Se $a^T x \geq \alpha$ é válida para $\mathcal{P}_X \implies \sum_{t_p \in P} \bar{f}_p \sum_{\ell \in A_\ell} a_\ell^p y_\ell^p \geq \alpha$ é válida para \mathcal{P}_Y .

Demonstração. Os itens (a), (b), (c) e (e) são consequências imediatas da definição dos poliedros (Definição 6.2). Resta provar o item (d).

Seja $a^T y \geq \alpha$ uma inequação válida para todo $y \in \mathcal{P}_Y$, logo

$$\sum_{\ell \in A_\ell} \sum_{t_p \in P} a_\ell^p y_\ell^p \geq \alpha.$$

Se tomarmos $a_\ell^* = \max_{t_p \in P} \{a_\ell^p\} > 0$, então temos que

$$\sum_{\ell \in A_\ell} a_\ell^* \sum_{t_p \in P} y_\ell^p \geq \sum_{\ell \in A_\ell} \sum_{t_p \in P} a_\ell^p y_\ell^p \geq \alpha. \quad (6.4)$$

Por outro lado, pela Definição 6.2, para todo $z \in \mathcal{P}_Z$, existe $\bar{y} \in \mathcal{P}_Y$, tal que

$$z_\ell \geq \sum_{t_p \in P} f_p \bar{y}_\ell^p, \forall \ell \in A_\ell.$$

Se tomarmos $f^* = \min_{t_p \in P} \{\bar{f}_p\}$, então temos que

$$z_\ell \geq \sum_{t_p \in P} f_p \bar{y}_\ell^p \geq f^* \sum_{t_p \in P} \bar{y}_\ell^p, \forall \ell \in A_\ell.$$

Ou ainda, como $f^* > 0$,

$$\frac{1}{f^*} z_\ell \geq \sum_{t_p \in P} \bar{y}_\ell^p, \forall \ell \in A_\ell. \quad (6.5)$$

Por hipótese a inequação (6.4) é válida para todo $y \in \mathcal{P}_Y$ e, em particular para \bar{y} , isto é,

$$\sum_{\ell \in A_\ell} a_\ell^* \sum_{t_p \in P} \bar{y}_\ell^p \geq \alpha. \quad (6.6)$$

Portanto, pode-se utilizar (6.5) para reescrever (6.6) como

$$\frac{1}{f^*} \sum_{\ell \in A_\ell} a_\ell^* z_\ell \geq \sum_{\ell \in A_\ell} a_\ell^* \sum_{t_p \in P} \bar{y}_\ell^p \geq \alpha.$$

Daí, conclui-se que $\sum_{\ell \in A_\ell} a_\ell^* z_\ell \geq \alpha f^*$. \square

6.1.1 Dimensão do Poliedro \mathcal{P}_S

Um dos aspectos fundamentais do estudo da estrutura facial de um poliedro está relacionado ao cálculo de sua dimensão, visto que tal informação permite, entre outras coisas, realizar uma caracterização/identificação das facetas do mesmo.

O número de variáveis de decisão (w_ℓ) é dado por $|A_\ell|$, enquanto que o número de variáveis de fluxo (f_ℓ^p) é $|A_\ell| \cdot |P|$. Por outro lado, para cada produto, existem $|N_\ell|$ restrições de conservação de fluxo das quais $|N_\ell| - 1$ são linearmente independentes. E, além disso, existem $|N_\ell^o|$ restrições de conservação/continuidade de comprimento de onda também linearmente independentes. Daí, um limite superior para a dimensão de \mathcal{P}_S , $\dim(\mathcal{P}_S)$, é dado por $|A_\ell| + |A_\ell| \cdot |P| - |P|(|N_\ell| - 1) - |N_\ell^o| = |P|(|A_\ell| - |N_\ell| + 1) + |A_\ell| - |N_\ell^o|$.

Em seguida, de modo a se provar que este limite é justo, isto é, que $\dim(\mathcal{P}_S) = |A_\ell|(|P| + 1) - |P|(|N_\ell| - 1) - |N_\ell^o|$ deveria se mostrar que não existem outras igualdades implícitas no modelo, além das já estabelecidas. Em outras palavras, se

$$\sum_{\ell \in A_\ell} \alpha_\ell w_\ell + \sum_{t_p \in P} \sum_{\ell \in A_\ell} \beta_\ell^p f_\ell^p = \gamma$$

for satisfeita por todas as soluções $(w, f) \in \mathcal{P}_S$, então essa equação é uma combinação linear das igualdades presentes no modelo (em verdade das $|P|(|N_\ell| - 1) + |N_\ell^o|$ igualdades linearmente independentes presentes no modelo).

Contudo, apesar dos esforços, não se conseguiu provar que o limite apresentado acima é justo, isto é, que $\dim(\mathcal{P}_S) = |A_\ell|(|P| + 1) - |P|(|N_\ell| - 1) - |N_\ell^o|$. As dificuldades encontradas nesse processo estão relacionadas diretamente com a forma de representação das capacidades de transporte (comprimentos de onda) através de camadas. Tal representação não parece permitir a obtenção de uma solução viável para o problema a partir de outra já existente (na verdade, não se conseguiu nem mesmo garantir a existência de uma outra solução). A prova mencionada acima se torna muito difícil (ou mesmo inviável) sem a possibilidade de se obter uma outra solução viável a partir de uma dada solução.

Em seguida, os poliedros da Definição 6.2 são descritos/examinados mais detalhadamente. O objetivo maior desse esforço é tentar obter no futuro inequações válidas para cada um deles e, aplicando-se os Lemas 6.3 e 6.1, inequações válidas para \mathcal{P}_M e \mathcal{P}_S , respectivamente.

6.1.2 Sobre o Poliedro \mathcal{P}_X

O poliedro \mathcal{P}_X representa a envoltória convexa do conjunto de soluções de um problema de fluxo não-capacitado de vários produtos, cuja dimensão é dada pela seguinte proposição.

Proposição 6.4. *A dimensão de \mathcal{P}_X , $\dim(\mathcal{P}_X)$, é dada por $|P|(|A_\ell| - |N_\ell| + 1)$.*

Demonstração. Como o número de variáveis é $|A_\ell||P|$ e, no conjunto de restrições (6.3b), existem apenas $|P|(|N_\ell| - 1)$ igualdades linearmente independentes, então um limite superior para a dimensão de \mathcal{P}_X será dado por $|A_\ell||P| - |P|(|N_\ell| - 1)$, ou melhor, $\dim(\mathcal{P}_X) \leq |P|(|A_\ell| - |N_\ell| + 1)$.

Em seguida, de modo a se provar que este limite é justo, isto é, que $\dim(\mathcal{P}_\mathcal{X}) = |P|(|A_\ell| - |N_\ell| + 1)$ deve-se mostrar que não existem outras igualdades implícitas além daquelas representadas pelas restrições (6.3b). Em outras palavras, se

$$\sum_{t_p \in P} \sum_{\ell \in A_\ell} \alpha_\ell^p x_\ell^p = \gamma \quad (6.7)$$

for satisfeita por todas as soluções $x \in \mathcal{P}_\mathcal{X}$, então essa equação é uma combinação linear das igualdades representadas pelas restrições (6.3b).

Primeiramente, mostrar-se-á que, para todo $t_p \in P$ e para todo ciclo \mathcal{C} em $G_\ell(N_\ell, A_\ell)$, $\sum_{\ell \in \mathcal{C}} \alpha_\ell^p = 0$. Como qualquer ciclo do grafo $G_\ell(N_\ell, A_\ell)$ pode ser decomposto em uma coleção de ciclos simples (isto é, sem repetição de nós), essa afirmação só precisa ser demonstrada para ciclos simples.

Considere uma solução qualquer $x \in \mathcal{P}_\mathcal{X}$ (por exemplo, aquela obtida a partir do envio do fluxo de cada produto $t_p \in P$ através do caminho mínimo entre sua origem s_p e seu destino d_p), pode-se, então, construir uma outra solução $\bar{x} \in \mathcal{P}_\mathcal{X}$, a partir de x , da seguinte forma. Para um dado produto $t_{\bar{p}} \in P$, selecione qualquer ciclo simples \mathcal{C} de G_ℓ . Pode-se aumentar o envio de fluxo do produto $t_{\bar{p}}$ através do ciclo \mathcal{C} sem violar nenhuma das restrições (6.3b) e (6.3g). Sendo assim, \bar{x} será construída de modo que $\bar{x}_\ell^p = x_\ell^p, \forall \ell \in A_\ell, \forall t_p \in P \setminus \{t_{\bar{p}}\}$. Já para o produto $t_{\bar{p}}$, temos $\bar{x}_\ell^{\bar{p}} = x_\ell^{\bar{p}}, \forall \ell \in A_\ell \setminus \mathcal{C}$ e $\bar{x}_\ell^{\bar{p}} = x_\ell^{\bar{p}} + 1, \forall \ell \in \mathcal{C}$. Como x e \bar{x} pertencem a $\mathcal{P}_\mathcal{X}$, então pode-se substituí-las em (6.7) e subtraindo membro a membro, obtém-se $\sum_{\ell \in \mathcal{C}} \alpha_\ell^{\bar{p}} = 0$, isto é, a somatória dos $\alpha_\ell^{\bar{p}}$ de um dado produto $t_{\bar{p}}$ ao longo dos arcos do ciclo \mathcal{C} é nula. Como $t_{\bar{p}}$ e \mathcal{C} foram escolhidos de forma arbitrária, conclui-se que $\forall t_p \in P, \sum_{\ell \in \mathcal{C}} \alpha_\ell^p = 0$, para todo ciclo \mathcal{C} .

Considere um nó qualquer $i \in N_\ell$ e um caminho $\bar{\mathcal{C}}$ de s_p a i em G_ℓ . Em seguida, mostrar-se-á que, para todo $t_p \in P$ e para todo $i \in N_\ell$, o valor de $\mu_i^p = \sum_{\ell \in \bar{\mathcal{C}}} \alpha_\ell^p$ é independente do caminho $\bar{\mathcal{C}}$ selecionado. Primeiramente, considerar-se-á o caso em que i seja um nó de *add-drop*, isto é, $i \in N_\ell^e$. A Proposição A.19 garante a existência de dois caminhos disjuntos $\bar{\mathcal{C}}_1$ e $\bar{\mathcal{C}}_2$ de s_p a $i \in N_\ell^e$. Além disso, a mesma proposição garante a existência de um caminho $\bar{\mathcal{C}}_3$ de i a s_p . Sendo assim, $\bar{\mathcal{C}}_1 \cup \bar{\mathcal{C}}_3$ forma um ciclo e, portanto, $\sum_{\ell \in \bar{\mathcal{C}}_1 \cup \bar{\mathcal{C}}_3} \alpha_\ell^p = \sum_{\ell \in \bar{\mathcal{C}}_1} \alpha_\ell^p + \sum_{\ell \in \bar{\mathcal{C}}_3} \alpha_\ell^p = 0$, ou ainda, $\sum_{\ell \in \bar{\mathcal{C}}_1} \alpha_\ell^p = -\sum_{\ell \in \bar{\mathcal{C}}_3} \alpha_\ell^p$. Por outro lado, $\bar{\mathcal{C}}_2 \cup \bar{\mathcal{C}}_3$ também forma um ciclo e, $\sum_{\ell \in \bar{\mathcal{C}}_2 \cup \bar{\mathcal{C}}_3} \alpha_\ell^p = \sum_{\ell \in \bar{\mathcal{C}}_2} \alpha_\ell^p + \sum_{\ell \in \bar{\mathcal{C}}_3} \alpha_\ell^p = 0$, ou ainda, $\sum_{\ell \in \bar{\mathcal{C}}_2} \alpha_\ell^p = -\sum_{\ell \in \bar{\mathcal{C}}_3} \alpha_\ell^p$. Logo, $\sum_{\ell \in \bar{\mathcal{C}}_1} \alpha_\ell^p = \sum_{\ell \in \bar{\mathcal{C}}_2} \alpha_\ell^p$, o que permite concluir que o caminho selecionado de s_p a $i \in N_\ell^e$ não afeta o valor de μ_i^p .

Caso i seja um nó ótico, isto é, $i \in N_\ell^o$, duas possibilidades devem ser analisadas. Na primeira, considerar-se-á que i seja um nó ótico de saída. Se $k \in N_\ell^e$ representar o nó de *add-drop* associado ao mesmo elemento de rede que i , logo existe o arco $(k, i) \in A_\ell$. Seja $\bar{\mathcal{C}}_1$ um caminho de s_p a i que não passa pelo arco (k, i) . Além disso, a Proposição A.19 garante a existência de um caminho $\bar{\mathcal{C}}_2$ de s_p a $k \in N_\ell^e$ (não necessariamente disjunto de $\bar{\mathcal{C}}_1$). Seja $\bar{\mathcal{C}}_3$ um caminho de k a s_p passando pelo arco (k, i) e $\bar{\mathcal{C}}_3^i$ a parte de tal caminho correspondente ao caminho de i a s_p , isto é, $\bar{\mathcal{C}}_3^i = \bar{\mathcal{C}}_3 \setminus \{(k, i)\}$. Sendo assim, tanto $\bar{\mathcal{C}}_2 \cup \{(k, i)\} \cup \bar{\mathcal{C}}_3^i$ como $\bar{\mathcal{C}}_1 \cup \bar{\mathcal{C}}_3^i$ formam ciclos e, portanto, $\sum_{\ell \in \bar{\mathcal{C}}_2 \cup \{(k, i)\} \cup \bar{\mathcal{C}}_3^i} \alpha_\ell^p = \sum_{\ell \in \bar{\mathcal{C}}_2} \alpha_\ell^p + \alpha_{ki}^p + \sum_{\ell \in \bar{\mathcal{C}}_3^i} \alpha_\ell^p = 0$ e $\sum_{\ell \in \bar{\mathcal{C}}_1 \cup \bar{\mathcal{C}}_3^i} \alpha_\ell^p = \sum_{\ell \in \bar{\mathcal{C}}_1} \alpha_\ell^p + \sum_{\ell \in \bar{\mathcal{C}}_3^i} \alpha_\ell^p = 0$. Daí, conclui-se $\sum_{\ell \in \bar{\mathcal{C}}_2 \cup \{(k, i)\}} \alpha_\ell^p = \sum_{\ell \in \bar{\mathcal{C}}_1} \alpha_\ell^p$, ou melhor, o caminho selecionado de s_p a um nó ótico de saída $i \in N_\ell^o$ não afeta o valor de μ_i^p .

Por outro lado, $i \in N_\ell^o$ pode ser um nó ótico de entrada. Seja $j \in N_\ell^o$ o nó ótico de saída para o qual existe o arco $(j, i) \in A_\ell$. A mesma linha de raciocínio anterior pode ser adotada para mostrar que existem dois caminhos entre s_p e o nó ótico de entrada i : $\bar{C}_2 \cup \{(k, j), (j, i)\}$ e $\bar{C}_1 \cup \{(j, i)\}$. Além disso, seja \bar{C}_3 um caminho de k a s_p passando pelo arcos (k, j) e (j, i) e \bar{C}_3^i a parte de tal caminho correspondente ao caminho de i a s_p , isto é, $\bar{C}_3^i = \bar{C}_3 \setminus \{(k, j), (j, i)\}$. A união de \bar{C}_3^i com qualquer um dos caminhos acima forma um ciclo e permite concluir que $\sum_{\ell \in \bar{C}_2 \cup \{(k, j), (j, i)\}} \alpha_\ell^p = \sum_{\ell \in \bar{C}_1 \cup \{(j, i)\}} \alpha_\ell^p$. Novamente, o caminho selecionado de s_p a um nó ótico de entrada $i \in N_\ell^o$ não afeta o valor de μ_i^p .

Finalmente, caso se multiplique cada uma das igualdades dadas pelas equações (6.3b) pelo correspondente valor de μ_i^p , obtém-se, a partir da soma, membro a membro, de todas elas a seguinte equação:

$$\sum_{t_p \in P} \sum_{i \in N_\ell} \mu_i^p \left(\sum_{\ell \in A_\ell^+(i)} x_\ell^p - \sum_{\ell \in A_\ell^-(i)} x_\ell^p \right) = \sum_{t_p \in P} \sum_{i \in N_\ell} \mu_i^p \bar{b}_i^p.$$

Rearranjando-se os termos acima, obtém-se:

$$\sum_{t_p \in P} \sum_{(i, j) \in A_\ell} (\mu_i^p - \mu_j^p) x_{ij}^p = \sum_{t_p \in P} \sum_{i \in N_\ell} \mu_i^p \bar{b}_i^p.$$

Para um dado conjunto de valores de μ_i^p , $\sum_{t_p \in P} \sum_{i \in N_\ell} \mu_i^p \bar{b}_i^p$ representa uma constante e fazendo-se $\alpha_{ij}^p = \mu_i^p - \mu_j^p, \forall (i, j) \in A_\ell$ resulta em:

$$\sum_{t_p \in P} \sum_{(i, j) \in A_\ell} \alpha_{ij}^p x_{ij}^p = \sum_{t_p \in P} \sum_{i \in N_\ell} \mu_i^p \bar{b}_i^p.$$

Isto demonstra que a igualdade (6.7) é realmente uma combinação linear das igualdades representadas pelas restrições (6.3b). \square

6.1.3 Sobre o Poliedro \mathcal{P}_y

O poliedro \mathcal{P}_y representa a envoltória convexa do conjunto de frações da demanda que são roteadas com sucesso em um problema de fluxo não-capacitado de vários produtos. Por sua vez, ele está relacionado ao problema de fluxo capacitado de vários produtos.

Definição 6.5. *Seja $\bar{\mathcal{Y}}$ o conjunto de soluções do problema de fluxo capacitado de vários produtos, em que \bar{y}_ℓ representa capacidade total a ser instalada no arco $\ell \in A_\ell$. Dessa forma:*

$$\bar{\mathcal{Y}} \equiv \left\{ \bar{y} \in \mathbb{R}^{|A_\ell|} \mid \exists x \in \mathcal{X} \text{ tal que } \sum_{t_p \in P} x_\ell^p \leq \bar{y}_\ell, \forall \ell \in A_\ell \right\}.$$

E associado a $\bar{\mathcal{Y}}$ temos o seguinte poliedro:

$$\mathcal{P}_{\bar{\mathcal{Y}}} \equiv \text{conv} \left\{ \bar{y} \in \mathbb{R}^{|A_\ell|} \mid \exists x \in \mathcal{X} \text{ tal que } \sum_{t_p \in P} x_\ell^p \leq \bar{y}_\ell, \forall \ell \in A_\ell \right\}.$$

Lema 6.6. Para os poliedros \mathcal{P}_Y , \mathcal{P}_Z e $\mathcal{P}_{\bar{Y}}$ pode-se afirmar que:

- (a) $\mathcal{P}_{\bar{Y}} \subseteq \mathcal{P}_Y$, e
- (b) $\mathcal{P}_Z \subseteq \mathcal{P}_{\bar{Y}}$.

Demonstração. Os item (a) e (b) são consequências da definição dos poliedros \mathcal{P}_Y e \mathcal{P}_Z (Definição 6.2) e do poliedro $\mathcal{P}_{\bar{Y}}$ (Definição 6.5) juntamente com o fato de que $x_\ell^p = \bar{f}_p y_\ell^p, \forall t_p \in P, \ell \in A_\ell$.

- (a) Seja $y \in \mathcal{Y}$, então $\bar{f}_p y \in \mathcal{X}$. Como $x_\ell^p = \bar{f}_p y_\ell^p, \forall t_p \in P, \ell \in A_\ell$, substituindo-se em $\sum_{t_p \in P} x_\ell^p \leq \bar{y}_\ell$ temos que:

$$\mathcal{P}_{\bar{Y}} \equiv \text{conv}\{\bar{y} \in \mathbb{R}^{|A_\ell|} \mid \exists y \in \mathcal{Y} \text{ tal que } \sum_{t_p \in P} \bar{f}_p y_\ell^p \leq \bar{y}_\ell, \forall \ell \in A_\ell\}.$$

Daí, pode-se concluir, imediatamente, que $\mathcal{P}_{\bar{Y}} \subseteq \mathcal{P}_Y$.

- (b) Seja $\bar{y} \in \bar{\mathcal{Y}}$, logo existe $y \in \mathcal{Y}$ tal que $\sum_{t_p \in P} \bar{f}_p y_\ell^p \leq \bar{y}_\ell, \forall \ell \in A_\ell$. Por outro lado, uma vez que $\bar{y} \in \mathbb{R}^{|A_\ell|}$ e $z \in \mathbb{Z}^{|A_\ell|}$, pode-se tomar $z_\ell = \lceil \bar{y}_\ell \rceil$, de modo que $z \in \mathcal{Z}$. Dessa forma, pode-se reescrever \mathcal{P}_Z como:

$$\mathcal{P}_Z \equiv \text{conv}\{z \in \mathbb{Z}^{|A_\ell|} \mid \exists \bar{y} \in \bar{\mathcal{Y}} \text{ tal que } \bar{y}_\ell \leq z_\ell, \forall \ell \in A_\ell\}.$$

Daí, conclui-se que $\mathcal{P}_Z \subseteq \mathcal{P}_{\bar{Y}}$. □

Considere, ainda, a definição do seguinte poliedro associado a relaxação linear do conjunto de soluções \mathcal{Z} .

Definição 6.7. Seja o poliedro \mathcal{LP}_Z o poliedro associado a relaxação linear do conjunto de soluções \mathcal{Z} , isto é,

$$\mathcal{LP}_Z \equiv \text{conv}\{z \in \mathbb{R}^{|A_\ell|} \mid \exists y \in \mathcal{Y} \text{ e } (y, z) \text{ satisfaz a (6.3d)}\}.$$

Dessa forma, é fácil verificar que $\mathcal{P}_{\bar{Y}}$ corresponde exatamente a \mathcal{LP}_Z , ou ainda, $\mathcal{P}_{\bar{Y}} \equiv \mathcal{LP}_Z$.

O poliedro $\mathcal{P}_{\bar{Y}}$, em particular, corresponde a envoltória convexa das soluções de um problema de instalação de capacidades em uma rede sujeita ao fluxo de vários produtos em que se admite a possibilidade de se instalar nos arcos da rede capacidades arbitrárias e em quantidades contínuas. Ele pode ser considerado um problema contínuo de projeto de rede capacitada com vários produtos, que pode ser resolvido em tempo polinomial através do método do elipsóide [90] ou o algoritmo de pontos interiores [86], uma vez que ele pode ser formulado como um problema linear com um número polinomial de variáveis e restrições. Além disso, uma caracterização exata das condições para que um vetor de capacidades seja suficiente para acomodar um roteamento viável (e contínuo) de produtos é dada pelo seguinte teorema.

Teorema 6.8 (Iri [80]; Kakusho e Onaga [85]). *Um vetor de capacidades \bar{y} é viável para um problema contínuo de projeto de rede capacitada com vários produtos se e somente se, para todo vetor $\mu \geq 0$,*

$$\sum_{\ell \in A_\ell} \mu_\ell \bar{y}_\ell \geq \sum_{t_p \in P} \pi_p^\mu \bar{f}_p \quad (6.8)$$

em que o vetor $\mu \in \mathbb{R}_+^{|A_\ell|}$ corresponde a pesos não negativos associados a cada arco $\ell \in A_\ell$ e π_p^μ representa o tamanho do caminho mínimo em G_ℓ entre s_p e d_p obtido utilizando-se μ_ℓ como comprimento do arco $\ell, \forall \ell \in A_\ell$.

A necessidade dessas inequações se origina no fato de que a forma mais barata de ser rotear um fluxo de vários produtos, caso não exista nenhuma restrição de capacidade e apenas os custos μ_ℓ sejam dados, é rotear cada produto separadamente através de seu caminho mínimo em relação ao vetor de custos μ . Dessa forma, um limite inferior para $\mu^T \bar{y}$ é a somatória dos produtos entre cada valor de tráfego (\bar{f}_p) e o tamanho do caminho mínimo (π_p^μ) entre os nós de oferta (s_p) e de demanda (d_p). A suficiência de (6.8) é uma consequência do teorema da dualidade em programação linear.

Essa caracterização das capacidades viáveis foi apresentada pela primeira vez em [80, 85]. As inequações (6.8) são denominadas *inequações métricas*. Esse nome é motivado pelo fato de que, para a situação em que G_ℓ é um grafo completo e em que exista tráfego entre cada um dos pares de nós, qualquer $\mu \in \mathbb{R}^{|A_\ell|}$ – definindo uma inequação não redundante em (6.8) – induz uma (*pseudo*-)métrica em G_ℓ , isto é, μ é não negativo, simétrico e atende a desigualdade triangular $\mu_{uv} + \mu_{vw} \geq \mu_{uw}$, para quaisquer nós $u, v, w \in N_\ell$, assumindo-se que $\mu_{uu} = 0$.

Podemos restringir nossa atenção em (6.8) às inequações definidas pelos vetores (μ, π) do conjunto de raios extremais do cone $\{\mu \in \mathbb{R}^{|A_\ell|}, \pi \in \mathbb{R}^{|P|} \mid \mu \geq 0, \pi_p = \pi_p^\mu, \forall t_p \in P\}$. Os raios extremais desse cone foram investigados em [4, 99].

Alguns casos de inequações métricas são muito importantes. Suponha que o conjunto de nós N_ℓ do grafo G_ℓ foi particionado em k conjuntos disjuntos N_1, \dots, N_k , de modo que $N_1 \cup \dots \cup N_k = N_\ell$ e $N_i \cap N_j = \emptyset, 1 \leq i < j \leq k$. Dados dois subconjuntos $N_1, N_2 \subseteq N_\ell$, em que $N_1 \cap N_2 = \emptyset$, o conjunto $\delta_{G_\ell}(N_1, N_2) = \{\ell \in A_\ell \mid \ell = (u, v), u \in N_1, v \in N_2\}$ contém todos os arcos com origem em um nó de N_1 destinados a um nó de N_2 . Considere que a definição de δ_{G_ℓ} para partição N_1, \dots, N_k é dada por

$$\delta_{G_\ell}(N_1, \dots, N_k) = \bigcup_{\substack{1 \leq i \leq k \\ 1 \leq j \leq k \\ i \neq j}} \delta_{G_\ell}(N_i, N_j)$$

e que adotou-se os seguintes valores como peso para os arcos $\ell \in A_\ell$:

$$\mu_\ell = \begin{cases} 1 & , \text{ se } \ell \in \delta_{G_\ell}(N_1, \dots, N_k), \\ 0 & , \text{ caso contrário.} \end{cases}$$

Dessa forma, obtém-se a seguinte inequação a partir de (6.8):

$$\sum_{\ell \in \delta_{G_\ell}(N_1, \dots, N_k)} \bar{y}_\ell \geq \sum_{t_p \in \delta_P(N_1, \dots, N_k)} \bar{f}_p \quad (6.9)$$

em que $\delta_P(N_1, \dots, N_k)$ representa o conjunto de produtos cujos nós de origem e de destino pertencem a subconjuntos distintos da partição N_1, N_2, \dots, N_k .

Para um $k > 2$ as inequações (6.9) são denominadas inequações de k -partição (*k-graph-partition inequality*), enquanto que para $k = 2$, elas são chamadas de inequações de corte (*cut inequality*). Vários pesquisadores têm investigado sob quais condições um vetor de capacidades é viável se e somente se todas as inequações de corte (ou as inequações de k -partição, $k \leq l$, para algum $l \in \mathbb{N}$ e fixo) são atendidas. Nesse sentido, dois resultados bem conhecidos são o Teorema Fluxo Máximo-Corte Mínimo e sua extensão para dois produtos.

Teorema 6.9 (Ford e Fulkerson [51]). *Para $|P| = 1$, um vetor de capacidades \bar{y} é viável para o problema contínuo de projeto de rede capacitada se e somente se \bar{y} satisfaz todas as inequações de corte (isto é, as inequações (6.9) com $k = 2$).*

Teorema 6.10 (Hu [75]). *Para $|P| = 2$, um vetor de capacidades \bar{y} é viável para o problema contínuo de projeto de rede capacitada se e somente se \bar{y} satisfaz todas as inequações de corte (isto é, as inequações (6.9) com $k = 2$).*

As inequações (6.9) são conhecidas por sua capacidade em promover melhorias nos valores obtidos por relaxações lineares de problemas de projeto e planejamento de redes [2, 21, 100]. Contudo, o problema de separação associado a elas é \mathcal{NP} -difícil.

Tais inequações já foram utilizadas previamente durante o cálculo de um dos limites inferiores de TGP (ver subseção 5.1.2) em que se particionou o conjunto de nós em dois subconjuntos: $N_1 = \{i\}$ e $N_2 = N_\ell \setminus N_1$, para cada nó de *add-drop* $i \in N_\ell^e$, e vice-versa¹.

6.1.4 Sobre o Poliedro \mathcal{P}_Z

O poliedro \mathcal{P}_Z representa a envoltória convexa das soluções de um problema de instalação de capacidades em uma rede sujeita ao fluxo de vários produtos em que se admite somente a possibilidade de se instalar nos arcos da rede capacidades em quantidades inteiras.

Em lugar de se examinar diretamente o poliedro \mathcal{P}_Z , a estrutura do poliedro associado ao problema de instalação de uma capacidade inteira em um único arco da rede será investigada. Para tanto, deve-se introduzir o seguinte poliedro.

Definição 6.11. *Seja Z_ℓ o conjunto de soluções do problema de instalação de uma capacidade inteira em um arco ℓ de uma dada rede, em que z_ℓ representa capacidade total (inteira) a ser instalada no arco ℓ . Dessa forma:*

$$Z_\ell \equiv \left\{ z_\ell \in \mathbb{Z} \mid \exists y_\ell \in \mathbb{R}^{|P|} \text{ tal que } \sum_{t_p \in P} \bar{f}_p y_\ell^p \leq z_\ell, 0 \leq y \leq 1 \right\}.$$

E associado a Z_ℓ temos o seguinte poliedro:

$$\mathcal{P}_{Z_\ell} \equiv \text{conv} \left\{ z_\ell \in \mathbb{Z} \mid \exists y_\ell \in \mathbb{R}^{|P|} \text{ tal que } \sum_{t_p \in P} \bar{f}_p y_\ell^p \leq z_\ell, 0 \leq y \leq 1 \right\}.$$

¹Em [20] tais inequações são denominadas inequações de grau – *degree inequations*

A estrutura facial do poliedro \mathcal{P}_{Z_ℓ} foi investigada em [101]. Neste trabalho, introduziu-se uma classe de inequações – inequações de capacidade residual – que juntamente com as restrições associadas a relaxação linear de Z_ℓ são suficientes para se descrever \mathcal{P}_{Z_ℓ} . Contudo, somente em [3], apresentou-se um algoritmo exato (e polinomial) de separação para tais inequações.

Seja $Q \subseteq P$ um subconjunto dos produtos, $\eta = \lceil \sum_{t_p \in Q} \bar{f}_p \rceil$ e

$$r = \begin{cases} \sum_{t_p \in Q} \bar{f}_p - \lfloor \sum_{t_p \in Q} \bar{f}_p \rfloor, & \text{caso } \eta \notin \mathbb{Z}; \\ 1, & \text{caso contrário.} \end{cases}$$

Então as inequações de capacidade residual [3, 101, 102] podem ser expressas da seguinte forma:

$$\sum_{t_p \in Q} \bar{f}_p (1 - y_\ell^p) \geq r(\eta - z_\ell). \quad (6.10)$$

Contudo, como $\bar{f}_p = \delta_p f_p \in \mathbb{Z}, \forall t_p \in P$, logo $\eta = \sum_{t_p \in Q} \bar{f}_p \in \mathbb{Z}$ e $r = 1$. Sendo assim, as inequações (6.10) podem ser reescritas como:

$$\sum_{t_p \in Q} \bar{f}_p (1 - y_\ell^p) \geq \eta - z_\ell \quad (6.11)$$

ou ainda:

$$\sum_{t_p \in Q} \bar{f}_p y_\ell^p - z_\ell \leq \sum_{t_p \in Q} \bar{f}_p - \eta \quad (6.12)$$

e finalmente:

$$\sum_{t_p \in Q} \bar{f}_p y_\ell^p \leq z_\ell. \quad (6.13)$$

Essas inequações implicam em termos da notação utilizada em $TGPC$ nas seguintes inequações:

$$\sum_{t_p \in Q} \delta_p f_\ell^p \leq \Delta w_\ell \quad (6.14)$$

que, por sua vez, se reduzem as inequações (3.10d) presentes na formulação $TGPC$ quando $Q = P$. Em outras palavras as restrições (3.10d) representam, na realidade, um subconjunto das inequações de capacidade residual associadas a essa formulação. Sob determinadas condições, pode-se mostrar que essas inequações também definem facetas para \mathcal{P}_{Z_ℓ} (ver Teorema 2 em [102]).

6.1.5 Considerações sobre o Estudo dos Poliedros

O estudo da estrutura facial do poliedro \mathcal{P}_S através de suas subestruturas e dos poliedros relacionados ($\mathcal{P}_X, \mathcal{P}_Y, \mathcal{P}_Z$ e \mathcal{P}_W) representa uma área promissora de pesquisa. Além disso, estudos recentes têm demonstrado que as inequações obtidas a partir desse tipo de investigação geralmente são muito úteis no desenvolvimento de métodos de *planos-de-corte* eficazes.

6.1.6 Inequações Válidas para \mathcal{P}_S

Em vista do que foi discutido na seção anterior, torna-se necessário uma avaliação da qualidade dos resultados obtidos através da abordagem poliédrica. Para tanto, optou-se por utilizar inicialmente três classes de inequações que são descritas a seguir.

Inequações de Corte de Capacidade

Considere dois subconjuntos N_1 e N_2 disjuntos de nós do grafo G_ℓ . Então, as inequações (6.9) se reduzem a seguinte forma:

$$\sum_{\ell \in \delta_{G_\ell}(N_1, N_2)} \bar{y}_\ell \geq \sum_{t_p \in \delta_P(N_1, N_2)} \bar{f}_p. \quad (6.15)$$

Como $\mathcal{P}_Z \subseteq \mathcal{P}_{\bar{y}}$ (ver prova do Lema 6.6 item (b)), então $z_\ell \geq \bar{y}_\ell$. Daí, se obtém que:

$$\sum_{\ell \in \delta_{G_\ell}(N_1, N_2)} z_\ell \geq \sum_{t_p \in \delta_P(N_1, N_2)} \bar{f}_p \quad (6.16)$$

ou ainda:

$$\sum_{\ell \in \delta_{G_\ell}(N_1, N_2)} \Delta w_\ell \geq \sum_{t_p \in \delta_P(N_1, N_2)} \delta_p f_p \quad (6.17)$$

e, como $\Delta = \delta_p m_p$, finalmente:

$$\sum_{\ell \in \delta_{G_\ell}(N_1, N_2)} w_\ell \geq \sum_{t_p \in \delta_P(N_1, N_2)} \frac{f_p}{m_p}. \quad (6.18)$$

O lado esquerdo das inequações (6.18) é inteiro, logo pode-se reescrever tais inequações da seguinte forma:

$$\sum_{\ell \in \delta_{G_\ell}(N_1, N_2)} w_\ell \geq \left\lceil \sum_{t_p \in \delta_P(N_1, N_2)} \frac{f_p}{m_p} \right\rceil. \quad (6.19)$$

Conforme mencionado anteriormente, um tipo especial dessas inequações (inequações de grau) já foi utilizado previamente, em que se particionou o conjunto de nós em dois subconjuntos: $N_1 = \{i\}$ e $N_2 = N_\ell \setminus N_1$, para cada nó de *add-drop* $i \in N_\ell^e$, e vice-versa (ver o cálculo de um dos limites inferiores de **TGP** na subseção 5.1.2).

A validade de tais inequações é facilmente constatada e inequações similares foram utilizadas, por exemplo, em [11, 20, 102]. Segundo [11], é possível mostrar que estas inequações definem facetas para $\mathcal{P}_{\bar{y}}$, quando os subgrafos induzidos por N_1 e N_2 são fortemente conexos e o somatório do lado direito de (6.19) não é inteiro. Mais ainda, segundo [20], apesar de sua aparente simplicidade as inequações de corte de capacidade são muito efetivas e geralmente violadas em grande número. Entretanto, elas podem não ser suficientes para se resolver qualquer instância, sendo, portanto, necessária a utilização de outras inequações descritas a seguir.

Inequações de Corte de Fluxo

As inequações de corte de fluxo podem ser vistas como uma generalização das inequações de corte de capacidade nas quais se incluem tanto as variáveis de decisão como as variáveis de fluxo.

Considere um subconjunto N_1 de nós do grafo G_ℓ e $Q \subseteq \delta_P(N_1, N_\ell \setminus N_1)$ um subconjunto dos produtos cujos nós de origem pertençam a N_1 e os nós de destino não pertençam. Seja $\delta_{G_\ell}^+(N_1)$ o conjunto de arcos cujos nós de origem pertencem a N_1 , isto é, $\delta_{G_\ell}^+(N_1) = \delta_{G_\ell}(N_1, N_\ell \setminus N_1)$. Além disso, suponha que $\delta_{G_\ell}^+(N_1)$ possa ser particionado em dois subconjuntos disjuntos A_1 e A_2 . Então, para a formulação TGP_M , as inequações de corte de fluxo possuem a seguinte forma:

$$\sum_{t_p \in Q} \sum_{\ell \in \delta_{G_\ell}^+(N_1)} x_\ell^p = \sum_{t_p \in Q} \sum_{\ell \in A_1} x_\ell^p + \sum_{t_p \in Q} \sum_{\ell \in A_2} x_\ell^p \geq \sum_{t_p \in Q} \delta_p f_p. \quad (6.20)$$

A validade de tais inequações também é facilmente constatada, visto que garantem que a quantidade total de fluxo dos produtos pertencentes a Q através dos arcos que saem de N_1 deve ser suficiente para atender a demanda desses produtos. Inequações similares podem ser encontradas em [20, 21].

Como a capacidade total instalada nos arcos pertencentes a A_2 deve ser maior que o somatório de fluxos através dos mesmos, pode-se reescrever (6.20) como:

$$\sum_{t_p \in Q} \sum_{\ell \in A_1} x_\ell^p + \sum_{\ell \in A_2} \Delta w_\ell \geq \sum_{t_p \in Q} \delta_p f_p \quad (6.21)$$

ou ainda, dividindo-se por Δ :

$$\sum_{t_p \in Q} \sum_{\ell \in A_1} x_\ell^p \geq \Delta \left(\sum_{t_p \in Q} \frac{f_p}{m_p} - \sum_{\ell \in A_2} w_\ell \right) \quad (6.22)$$

e, substituindo-se x_ℓ^p por $\delta_p f_\ell^p$, finalmente:

$$\sum_{t_p \in Q} \sum_{\ell \in A_1} \delta_p f_\ell^p \geq \Delta \left(\sum_{t_p \in Q} \frac{f_p}{m_p} - \sum_{\ell \in A_2} w_\ell \right). \quad (6.23)$$

Novamente, como o lado esquerdo das inequações (6.23) é inteiro, logo pode-se reescrever tais inequações da seguinte forma:

$$\sum_{t_p \in Q} \sum_{\ell \in A_1} \delta_p f_\ell^p \geq \left\lceil \Delta \left(\sum_{t_p \in Q} \frac{f_p}{m_p} - \sum_{\ell \in A_2} w_\ell \right) \right\rceil. \quad (6.24)$$

Inequações de 3-Partição

Finalmente, considere uma partição do conjunto de nós do grafo G_ℓ em três subconjuntos disjuntos N_1 , N_2 e N_3 . Semelhante ao que foi realizado em [20, 102], pode-se obter as seguintes inequações de 3-partição (ver inequações (6.9)):

$$\sum_{\ell \in \delta_{G_\ell}(N_1, N_2)} \bar{y}_\ell + \sum_{\ell \in \delta_{G_\ell}(N_1, N_3)} \bar{y}_\ell + \sum_{\ell \in \delta_{G_\ell}(N_2, N_3)} \bar{y}_\ell \geq \sum_{t_p \in \delta_P(N_1, N_2)} \bar{f}_p + \sum_{t_p \in \delta_P(N_1, N_3)} \bar{f}_p + \sum_{t_p \in \delta_P(N_2, N_3)} \bar{f}_p. \quad (6.25)$$

Essas inequações podem ser reescritas de forma análoga ao que foi realizado para inequações de corte de capacidade. Sendo assim, obtém-se as seguintes inequações:

$$\sum_{\ell \in \delta_{G_\ell}(N_1, N_2)} w_\ell + \sum_{\ell \in \delta_{G_\ell}(N_1, N_3)} w_\ell + \sum_{\ell \in \delta_{G_\ell}(N_2, N_3)} w_\ell \geq \sum_{t_p \in \delta_P(N_1, N_2)} \frac{f_p}{m_p} + \sum_{t_p \in \delta_P(N_1, N_3)} \frac{f_p}{m_p} + \sum_{t_p \in \delta_P(N_2, N_3)} \frac{f_p}{m_p}. \quad (6.26)$$

Novamente, como o lado esquerdo das inequações (6.26) é inteiro, logo pode-se reescrever tais inequações da seguinte forma:

$$\sum_{\ell \in \delta_{G_\ell}(N_1, N_2)} w_\ell + \sum_{\ell \in \delta_{G_\ell}(N_1, N_3)} w_\ell + \sum_{\ell \in \delta_{G_\ell}(N_2, N_3)} w_\ell \geq \left\lceil \sum_{t_p \in \delta_P(N_1, N_2)} \frac{f_p}{m_p} + \sum_{t_p \in \delta_P(N_1, N_3)} \frac{f_p}{m_p} + \sum_{t_p \in \delta_P(N_2, N_3)} \frac{f_p}{m_p} \right\rceil. \quad (6.27)$$

Na verdade, a partir das inequações (6.9), pode-se obter uma família de inequações de k -partição, bastando se fixar $k > 2$, isto é, utilizar uma partição do conjunto de nós do grafo G_ℓ em k subconjuntos disjuntos.

6.2 Avaliação Preliminar dos Cortes

Nesta seção, apresenta-se uma avaliação preliminar da utilização dos *planos-de-corte* dados pelas inequações da seção anterior na resolução do TGP. Para tanto, implementou-se “parcialmente” um método de *planos-de-corte* em C++ utilizando a biblioteca de funções do COIN-CBC [78]. Nesta implementação não se procurou resolver por completo o TGP e ela sequer possuía uma heurística para geração de limites superiores.

A principal motivação de tal esforço era obter um conjunto de resultados preliminares capazes de ajudar na condução e bom andamento deste trabalho. Sendo assim, essa implementação limita-se a gerar (e avaliar) uma série de *planos-de-corte* a partir de uma versão inicial simplificada do problema². Esse problema inicial só possuía as restrições de conservação de *containers* (ou de continuidade dos *lightpaths* – restrições (3.10c)), já as restrições de integralidade das variáveis de decisão foram relaxadas e todas as variáveis de fluxo foram removidas.

Um vez resolvido o problema inicial, a cada iteração do algoritmo implementado, uma série de *planos-de-corte* era gerada e acrescentada à formulação que, por sua vez, era resolvida novamente. Nos testes apresentados abaixo utilizaram-se as inequações de corte de capacidade e de 3-partição descritas anteriormente, juntamente com as inequações de grau (adicionadas à formulação apenas durante a primeira iteração). Para resolução dos problemas de programação linear gerados durante esse processo utilizou-se o COIN-CLP [78].

Durante a obtenção dos *planos-de-corte* torna-se necessário a resolução dos problemas de separação associados a cada conjunto de inequações, isto é, encontrar inequações válidas que sejam violadas pela solução atual. Contudo, como tais problemas são \mathcal{NP} -difíceis³, na implementação testada optou-se pela utilização de uma heurística simples e extremamente eficiente.

²A adoção de uma versão inicial simplificada (com um número menor de restrições) é um artifício comum na literatura da área, sendo utilizado, via de regra, quando se encontram dificuldades na obtenção de uma solução para relaxação linear do próprio problema original.

³O problema de separação das inequações de corte de capacidade se reduz ao problema do corte máximo que é \mathcal{NP} -difícil [11].

Tabela 6.1: Resultado dos *Cortes* nas Redes em Anel (valores médios)

$ V $	$ N $	$ A $	$ P $	Var	Restr	Cortes	Tempo (seg)	Gap (%)
4	20	32	40	32	16	40	0,02	0,00
			80	32	16	40	0,04	0,00
			120	32	16	40	0,06	3,91
5	25	40	40	40	20	58	0,03	0,00
			80	40	20	58	0,05	0,00
			120	40	20	67	0,09	3,43
6	30	48	40	48	24	84	0,05	0,00
			80	48	24	84	0,07	0,00
			120	48	24	88	0,12	0,43
9	45	72	40	72	36	178	0,10	0,22
			80	72	36	178	0,16	0,00
			120	72	36	189	0,27	0,84
10	50	80	40	80	40	220	0,13	0,00
			80	80	40	220	0,19	0,00
			120	80	40	220	0,27	0,52
16	80	128	40	128	64	541	0,49	3,76
			80	128	64	544	0,65	0,00
			120	128	64	544	0,81	2,79

A seleção dos subconjuntos para os quais a validade das inequações era verificada foi feita randomicamente. No caso das inequações de corte de capacidade, os nós de *add-drop* da rede foram repartidos aleatoriamente em quantidades iguais entre os dois subconjuntos (N_1 e N_2). Já nas inequações de 3-partição, o processo de distribuição aleatória dos nós de *add-drop* entre os 3 subconjuntos aconteceu livremente desde que se garantisse que nenhum subconjunto ficasse vazio. De modo a aumentar a probabilidade de se encontrar algumas inequações válidas violadas pela solução atual, o processo de geração randômica de subconjuntos foi repetido várias vezes a cada iteração ($n_e^2/2$ tentativas por iteração).

A implementação descrita acima foi testada nos mesmos problemas apresentados na seção 5.6, utilizando-se da mesma função de custo. Além disso, o número máximo de iterações foi limitado a 100 e se estabeleceu o *gap de dualidade* mínimo em 0,5%. Caso nenhuma nova inequação válida violada fosse encontrada em determinada iteração, o algoritmo também era interrompido.

Para o conjunto I (90 instâncias do **TGP** para redes em anel), os resultados obtidos pela implementação parcial do método de *planos-de-corte* (em um processador Pentium 4 de 1,8 GHz com 512 MB RAM) são mostrados na Tabela 6.1. A primeira coluna desta tabela apresenta o número de nós na rede original; enquanto que, na segunda e terceira colunas são apresentados, respectivamente, o número de nós e o número de arcos da representação estendida associada a cada uma das redes. A quarta coluna exhibe o número de produtos de cada subconjunto. O número de variáveis e o número de restrições da formulação inicial simplificada são mostrados na quinta e sexta colunas. O número médio de cortes gerados e o tempo médio gasto pelo algoritmo para cada subconjunto são apresentados nas duas colunas seguintes,

Tabela 6.2: Resultado dos *Cortes* nas Redes Irregulares (valores médios)

$ V $	$ E $	$ N $	$ A $	$ P $	Var	Restr	Cortes	Tempo (seg)	Gap (%)
6	8	38	76	40	76	32	84	0,06	0,00
				80	76	32	84	0,10	0,00
				120	76	32	84	0,16	0,15
10	24	58	108	40	108	48	220	0,16	0,00
				80	108	48	220	0,24	0,00
				120	108	48	225	0,42	0,43
15	21	99	230	80	230	84	478	0,92	0,07
				120	230	84	478	1,12	0,26
20	48	116	216	80	216	96	840	1,59	0,32
				120	216	96	840	1,89	0,14

seguidos pelo *gap de dualidade* médio calculado utilizando-se como limite superior a solução ótima obtida pelo *CPLEX* (durante os testes apresentados na seção 5.6) ou o melhor valor de limite superior obtido pela heurística lagrangeana (para os casos em que o *CPLEX* não foi capaz de encontrar uma solução ótima).

Como pode ser facilmente constatado, apesar da seleção randômica de cortes, a implementação exibe um desempenho excepcional, sendo seus resultados muito promissores. Entretanto, o mesmo fenômeno observado durante os testes apresentados na seção 5.6 parece ocorrer novamente, isto é, dificuldade em solucionar algumas instâncias pequenas e sobrecarregadas, bem com algumas instâncias grandes com pouco volume de tráfego. Tal situação parece estar relacionada ao conjunto de inequações válidas utilizadas e a “real” existência de um *gap de dualidade*.

O conjunto II contém 50 instâncias do **TGP** para redes irregulares e, na Tabela 6.2, apresentam-se os resultados para esse conjunto. Ela possui as mesmas colunas da Tabela 6.1, exceto pela inserção da segunda coluna em que o número de arcos da topologia irregular original é exibido. Novamente, os resultados obtidos são muito satisfatórios.

Já a Tabela 6.3 apresenta os resultados detalhados para cada uma dessas instâncias do conjunto III (10 instâncias baseadas na rede Pan-Européia da Fig. 5.4), além dos valores médios para cada subconjunto.

Finalmente, a Tabela 6.4 exibe os resultados obtidos pela aplicação do algoritmo à rede em anel com 4 nós descrita na seção 5.7. Para tanto foi utilizada a mesma matriz de tráfego (Tabela 5.12); variando-se a taxa de *grooming*, isto é, a capacidade máxima de transporte de um comprimento de onda (ou *container virtual*) entre os valores 3, 12 e 48. Os resultados obtidos são bem semelhantes aos apresentados nas tabelas anteriores. Deve-se destacar que ao se utilizar uma taxa de *grooming* igual a 12, o algoritmo foi capaz de obter um limite inferior igual ao valor fornecido para a solução ótima do problema. Contudo, na medida em que a capacidade de transporte é reduzida (ou que o volume do tráfego aumenta), o mesmo não ocorreu (por exemplo, para o caso em que a taxa de *grooming* é igual a 3).

Em primeiro lugar, deve-se lembrar que o valor do limite superior utilizado neste caso representa a melhor solução viável obtida (não sendo necessariamente a solução ótima do problema). Este fato talvez fosse o bastante para justificar o elevado *gap*

Tabela 6.3: Resultado dos *Cortes* na Rede Pan-Européia da Fig. 5.4

$ V $	$ E $	$ N $	$ A $	$ P $	Var	Restr	Núm. da Instância	Cortes	Tempo (seg)	Gap (%)
20	78	176	528	80	528	156	1	840	4,03	0,87
							2	840	4,10	0,65
							3	840	4,11	0,06
							4	839	3,98	0,71
							5	840	4,04	0,71
				Média	840	4,05	0,60			
				120	528	156	1	840	4,56	0,44
							2	840	4,59	4,86
							3	840	3,67	0,60
							4	840	4,60	0,11
5	840	4,50	0,97							
Média	840	4,38	1,40							

de dualidade apresentado neste caso. Entretanto, como observado em [20], apesar das inequações utilizadas serem extremamente simples e, geralmente, violadas em grandes quantidades, elas podem não ser suficientes para se solucionar um problema, especialmente quando o volume de tráfego é extremamente elevado (ou pequeno) se comparado a capacidade total disponível. Tal fenômeno necessita ser melhor explorado, de modo a se obter, eventualmente, inequações que ajudem a contornar essas dificuldades.

6.3 Método de *Planos-de-Corte*

Conforme descrito na subseção 4.2.2, um método de *planos-de-corte* para resolução do **TGP** foi implementado e testado. Contudo, diante das dificuldades de se obter uma “boa” descrição do fecho inteiro do poliedro associado às soluções viáveis do **TGP**, não se deve esperar que tal método seja capaz de encontrar os soluções ótimas para o problema.

A implementação desenvolvida inicia-se utilizando uma heurística para obtenção de um limite superior inicial (solução viável) para o problema. Na implementação, em questão, optou-se por utilizar a mesma heurística adotada nas abordagens la-grangeanas, um vez que ela, além de ser eficiente, apresentou um bom desempenho (isto é, conseguiu encontrar várias soluções ótimas).

Em seguida, o método de *planos-de-corte* realiza um procedimento iterativo da

Tabela 6.4: Resultado dos *Cortes* para o Anel da Seção 5.7

Taxa de <i>Grooming</i>	Var	Restr	Cortes	Tempo (seg.)	Gap (%)
3	160	80	70	0,08	4,95
12	64	32	41	0,02	0,34
48	32	16	40	0,02	0,00

seguinte forma. A cada iteração a relaxação linear da formulação TGP_C apresentada na subseção 3.1.4 acrescida de alguns novos cortes (inequações válidas) é resolvida. Além disso, a heurística para obtenção de soluções viáveis também é aplicada a cada iteração.

Além dos *planos-de-corte* avaliados na seção anterior (isto é, inequações de corte de capacidade e de k -partição), utilizou-se também na implementação dois tipos de cortes genéricos (*general purpose cutting planes*)⁴: **MIR** e *Probing* – ambos disponíveis na biblioteca de cortes do COIN-CGL [78]. O primeiro tipo, **MIR** – *Mixed Integer Rounding*, foi introduzido em [103] e utilizado com relativo sucesso na resolução de problema de programação inteira mista. Já o segundo tipo, *Probing*, apesar de não ser bem documentado, representa uma implementação simplificada das idéias sobre “sondagem” de variáveis (*probing*) utilizadas no OSL [160]. Em verdade, para se estabelecer esse tipo de corte investiga-se o efeito da redução (ou do aumento) de valor das variáveis inteiras presentes na formulação (ou de um subconjunto delas).

O método de *planos-de-corte* foi executado até que uma das seguintes condições fosse verdadeira:

- *gap de dualidade* inferior a 0,10%; ou
- não se conseguisse encontrar nenhuma inequação válida (corte) que fosse violada pela solução atual; ou
- não se conseguisse uma melhoria de pelo menos 0,10% no limite inferior; ou
- o número máximo de iterações fosse alcançado.

6.4 Algoritmo de *Branch-and-Cut*

O objetivo final da abordagem poliédrica do **TGP** era o desenvolvimento de um algoritmo de *branch-and-cut* para o problema. Conforme descrito anteriormente, essa implementação foi realizada utilizando-se da biblioteca de funções do COIN-CBC [78]. Apesar de todo o procedimento de busca ter sido realizado de forma transparente pelo conjunto de rotinas do COIN-CBC, deve-se destacar alguns detalhes importantes sobre a implementação realizada.

Em primeiro lugar, de modo a reduzir o esforço computacional, optou-se por resolver inicialmente a relaxação linear da formulação TGP_C em que se substituiu as restrições de conservação de fluxo – restrições (3.10b) – pelas inequações de grau. A solução obtida, então, foi utilizada para se iniciar a busca de uma solução para a relaxação linear da formulação TGP_C completa⁵. Esse procedimento demonstrou ser

⁴Esses cortes foram selecionados após a realização de diversos testes com os mais variados tipos de cortes.

⁵O COIN-CBC se mostrou incapaz de resolver uma versão simplificada da formulação do problema. Isto se deve ao fato dele utilizar uma heurística de “arredondamento” durante a busca. Dessa forma, caso se forneça uma formulação simplificada para o COIN-CBC (sem as restrições de conservação de fluxo, por exemplo), ele é capaz de resolver o problema “arredondando” o valores das variáveis de decisão (desconsiderando por completo as variáveis de fluxo).

muito eficaz, uma vez que a solução do problema sem as restrições de conservação de fluxo podia ser obtida muito rapidamente.

A geração de *planos-de-corte* foi realizada de forma semelhante àquela descrita para o método de *planos-de-corte*. Contudo, de modo a se reduzir o esforço computacional durante a busca, procurou-se gerar novas inequações de corte de capacidade e de k -partição a intervalos regulares (na implementação testada, a cada 10 nós explorados).

Já os cortes genéricos (**MIR** e *Probing*) só foram aplicados na exploração do nó inicial. Isto porque testes preliminares indicaram que nenhuma grande melhoria no processo de busca era obtida a partir da utilização desses cortes durante a exploração de outro nó da árvore (diferente da raiz), além do fato de que eles são computacionalmente caros. Além disso, a geração de cortes foi aplicada três vezes durante a exploração do nó inicial (problema original), enquanto que, durante a exploração dos demais nós da árvore, o procedimento de geração de cortes foi utilizado uma única vez.

Por fim, de modo a se agilizar a obtenção de novas soluções viáveis, além da heurística utilizada nas abordagens lagrangeanas, utilizou-se também uma heurística de “arredondamento” disponibilizada pela biblioteca de funções do COIN-CBC.

6.5 Algoritmo de *Cut-and-Branch*

Conforme descrito anteriormente, o algoritmo de *cut-and-branch* pode ser visto como uma implementação “simplificada” do processo de busca. A implementação realizada e testada neste trabalho segue as mesmas especificações do algoritmo de *branch-and-bound* descrito na seção anterior, exceto pelo fato de que o procedimento de geração de novos cortes só é utilizado durante a exploração do nó inicial.

6.6 Experimentos Computacionais

6.6.1 Resultados do Método de *Planos-de-Corte*

O método de *planos-de-corte* descrito anteriormente foi implementado em C++, utilizando a biblioteca de funções do COIN-CBC [78]. Os testes foram realizados na mesma plataforma utilizada pela avaliação preliminar (isto é, um processador Pentium 4 de 1,8 GHz com 512 MB RAM).

Os experimentos computacionais foram realizados com as mesmas instâncias descritas no capítulo anterior (150 problemas divididos em 3 conjuntos) e o número máximo de iterações foi fixado em 100. Infelizmente, devido ao número excessivo de variáveis e restrições, o método de *planos-de-corte* não conseguiu obter resultados satisfatórios quando aplicado às grandes instâncias do problema (redes em anel com 16 nós e redes irregulares com 15 e 20 nós). Em verdade, cada iteração em que se resolve a relaxação linear (acrescida dos cortes) passou a consumir mais de uma hora de processamento para essas instâncias. Esse fato se assemelha ao ocorrido durante os testes utilizando o *CPLEX* e acabou por inviabilizar a realização de testes com grandes instâncias. Maiores explicações sobre esse fato são fornecidas adiante.

Tabela 6.5: Resultados do Método de *Planos-de-Corte* para Redes em Anel (valores médios)

Instância	Iter	Tempo (seg)	Gap (%)	Total <i>Cuts</i>	<i>Tight</i> (%)	Número de Cortes			ADMs	LPs	WLs
						Part ^a	Prob	MIR			
C04-40	2	0,48	0,00	292,4	59,4	17,0	246,2	29,2	4,0	4,0	8,0
C04-80	2	2,01	0,00	580,8	53,8	127,8	418,2	34,8	4,0	4,0	8,0
C04-120	8	38,80	17,05	1614,0	38,2	163,2	757,2	693,6	5,6	5,2	10,4
C05-40	2	0,83	0,00	453,4	55,2	149,0	273,6	30,8	5,0	5,0	10,0
C05-80	2	3,39	0,00	588,6	36,6	151,8	393,0	43,8	5,0	5,0	10,0
C05-120	6	38,84	17,77	1587,0	42,8	45,8	859,8	681,4	6,8	6,6	13,6
C06-40	2	2,74	0,00	1845,4	68,5	1469,2	334,4	41,8	6,0	6,0	12,0
C06-80	2	6,59	0,00	727,2	44,5	19,8	646,2	61,2	6,0	6,0	12,0
C06-120	7	78,34	22,66	3886,2	28,6	2216,2	1032,2	637,8	9,0	8,2	16,4
C09-40	2	6,92	0,00	2214,0	68,0	1453,4	687,4	73,2	9,0	9,0	18,0
C09-80	2	9,40	0,00	695,0	60,9	0,2	631,8	63,0	9,0	9,0	18,0
C09-120	10	267,66	35,63	13494,6	48,1	10882,4	1821,4	790,8	15,6	15,0	30,0
C10-40	2	10,02	0,00	2635,4	58,6	1938,8	615,2	81,4	10,0	10,0	20,0
C10-80	2	17,86	0,00	1120,2	57,4	3,0	1023,6	93,6	10,0	10,0	20,0
C10-120	7	313,88	23,56	14914,8	65,4	12557,6	1641,4	715,8	15,0	14,2	28,4

^a Número de cortes de capacidade e 3-partição

A Tabela 6.5 exibe os resultados obtidos pelo método de *planos-de-corte* para as redes em anel. A primeira coluna identifica o subconjunto de instâncias. Os valores médios para o número de iterações, para o tempo gasto (em segundos) e para o *gap de dualidade* são apresentados na segunda, terceira e quarta colunas, respectivamente. As duas colunas seguintes exibem o número médio de cortes gerados e o percentual de cortes “justos” (ou *tight*) ao final da execução do método. Em seguida, são apresentados os valores médios para as quantidades de cortes gerados de cada um dos tipos. Finalmente, as três últimas colunas exibem os valores médios para o número de multiplexadores *add-drop*, para o número de *lightpaths* e para o número de *wavelength-links* das melhores soluções obtidas pelo método de *planos-de-corte*.

Supreendentemente, o método de *planos-de-corte* conseguiu obter a solução ótima de várias instâncias em intervalos de tempo razoáveis – o que serve como um indicador da “boa” qualidade dos cortes utilizados. Além disso, também é fácil de se constatar que, a medida que o tamanho da rede cresce juntamente com seu tráfego, o método de *planos-de-corte* encontrou maiores dificuldades em obter boas soluções (chegando a exibir valores médios para o *gap de dualidade* acima de 20%).

Deve-se, ainda, ressaltar que apesar do elevado número de cortes obtidos uma fração considerável deles (aproximadamente 50%) são “justos” (e ativos) na solução final. Por outro lado, isto pode indicar uma das razões que inviabilizaram a resolução de grandes instâncias, pois na medida que o tamanho da rede cresce (independente do volume de tráfego ser muito baixo ou muito alto), um número cada vez maior de cortes é gerado e acrescido à formulação. No caso das grandes instâncias do **TGP**, que já possuem um grande número de variáveis e restrições⁶, o acréscimo dessa grande quantidade de cortes pode, pelo menos em parte, justificar a dificuldade em se resolver a relaxação linear do problema.

Já a Tabela 6.6 apresenta os resultados obtidos pelo método de *planos-de-corte* para as redes irregulares. As colunas nessa tabela são as mesmas da tabela anterior.

⁶Por exemplo, para as redes em anel com 16 nós e 120 produtos são utilizadas 15360 variáveis e 9792 restrições na formulação original, enquanto que para as redes irregulares com 15 nós e 120 produtos a formulação original possui 27830 variáveis e 12294 restrições.

Tabela 6.6: Resultados do Método de *Planos-de-Corte* para Redes Irregulares (valores médios)

Instância	Iter	Tempo (seg)	Gap (%)	Total <i>Cuts</i>	<i>Tight</i> (%)	Número de Cortes			ADMs	LPs	WLs
						Part ^a	Prob	MIR			
M06-40	2	7,91	0,00	753,6	74,0	261,4	409,4	82,8	6,0	6,0	12,0
M06-80	3	34,91	0,00	987,8	57,0	261,2	642,0	84,6	6,0	6,0	12,0
M06-120	9	459,09	26,61	4094,6	35,2	1479,6	1774,4	840,6	9,4	8,6	17,2
M10-40	3	44,29	0,00	6704,8	83,1	5795,4	781,4	128,0	10,0	10,0	20,0
M10-80	3	107,85	0,00	7219,6	56,3	5803,0	1260,0	156,6	10,0	10,0	20,0
M10-120	7	907,79	17,93	21169,2	66,7	18379,2	2473,0	317,0	13,4	12,8	25,6

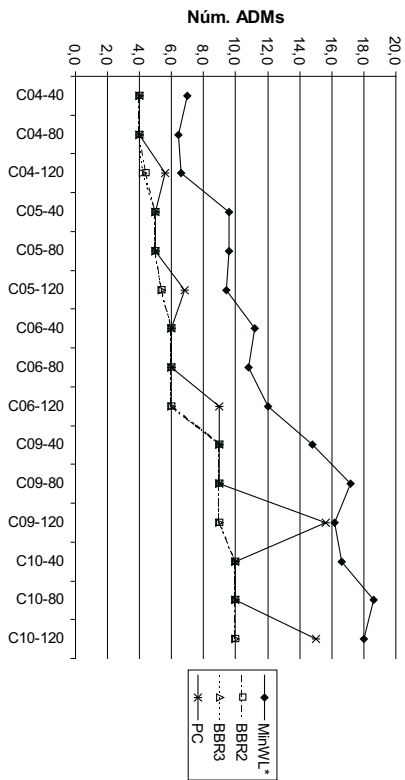
^a Número de cortes de capacidade e 3-partição

Um exame dos resultados apresentados confirma as mesmas observações realizadas para as redes em anel. Contudo, um elevado aumento nos valores médios de tempo gasto deve ser destacado. Novamente, o método de *planos-de-corte* foi capaz de obter a solução ótima para várias instâncias. No total (para redes em anel e irregulares), ele encontrou a solução ótima para 83 das 105 instâncias (79%).

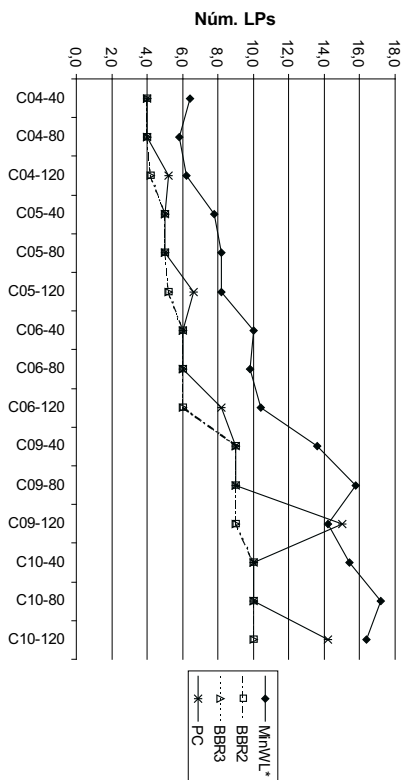
Os gráficos da Figura 6.1 apresentam uma comparação dos valores médios obtidos pelo método de *planos-de-corte* com aqueles obtidos (e apresentados no capítulo anterior) pela heurística proposta em [166] – utilizando a política **MinWL***, bem como com os resultados dos algoritmos de *branch-and-bound* **BBR2** e **BBR3** para as redes em anel. Nesses gráficos, os resultados obtidos pelo método de *planos-de-corte* estão indicados na legenda pela sigla **PC**. Deve-se mencionar que o método de *planos-de-corte* obteve resultados praticamente equivalentes aos obtidos pelos algoritmos **BBR2** e **BBR3** em relação ao número médio de multiplexadores de *add-drop* e de *lightpaths*. A exceção ficou por conta das instâncias com alto volume de tráfego (isto é, com 120 produtos), situação em que o método de *planos-de-corte* apresentou resultados tão ruins quanto os da heurística. Por fim, vale lembrar que, para esses casos, o método de *planos-de-corte* não conseguiu encontrar a solução ótima (todas as 22 instâncias em que não se obteve a solução ótima possuíam um tráfego de 120 produtos), o que serve para justificar os resultados obtidos, uma vez que as soluções encontradas por **BBR2** e **BBR3** são praticamente todas ótimas ou “quase-ótimas” para essas instâncias.

Examinando-se o gráfico da Figura 6.1(c), constata-se que o número médio de *wavelength-links* alocados pelo método de *planos-de-corte* é praticamente o mesmo que o utilizado pelas duas abordagens lagrangeanas. Já no gráfico da Figura 6.1(d), verifica-se que o número médio de “saltos” (*hops*) na topologia virtual é praticamente constante (igual a 2), sendo duas vezes maior que aquele obtido pela heurística. Esse comportamento é semelhante ao das abordagens lagrangeanas e conduz a mesma conclusão: a melhor utilização dos recursos de rede (obtendo soluções com um número menor de ADMs e *lightpaths*) parece estar ligada à utilização de um número maior de “saltos” (*hops*) na topologia virtual para rotear as requisições de tráfego.

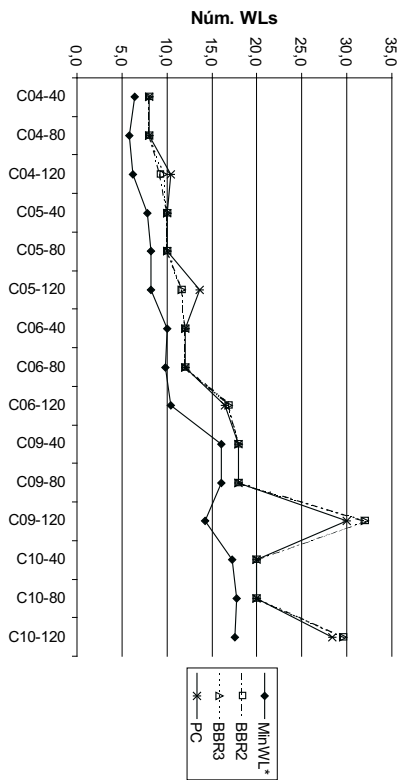
Já os gráficos da Figura 6.2 apresentam uma comparação dos valores médios obtidos pelo método de *planos-de-corte* com aqueles obtidos (e apresentados no capítulo anterior) pela heurística proposta em [166] – utilizando a política **MinWL***, bem como com os resultados dos algoritmos de *branch-and-bound* **BBR2** e **BBR3** para as redes irregulares. Novamente, os padrões observados para as redes em anel se repetem para as redes irregulares.



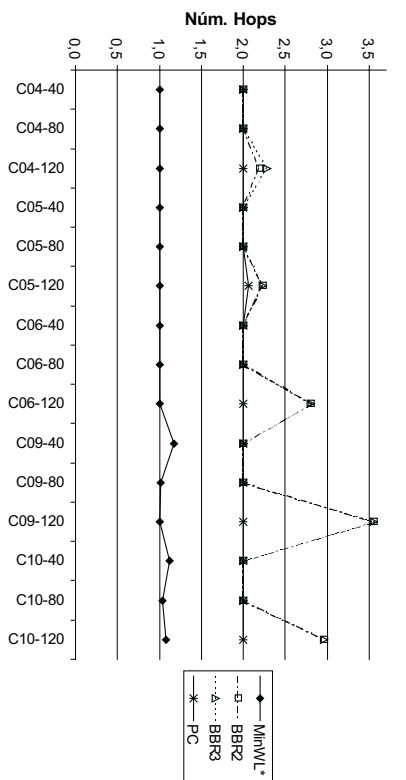
(a) Comparação do Número de ADMs



(b) Comparação do Número de LPs

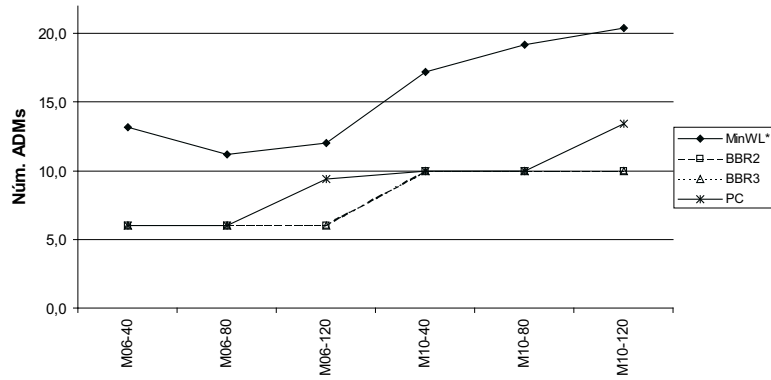


(c) Comparação do Número de WLS

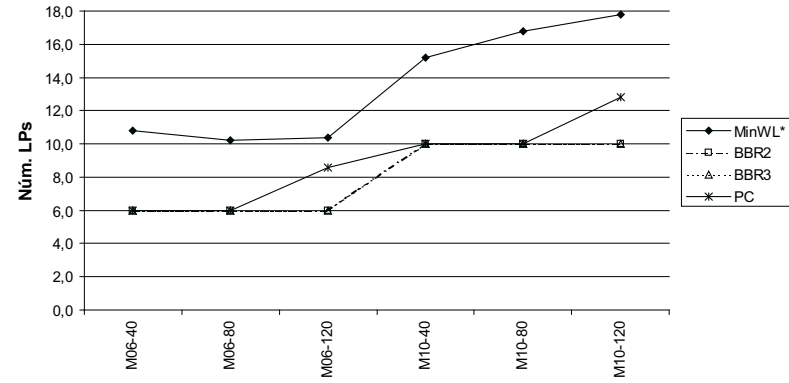


(d) Comparação do Número de Hops

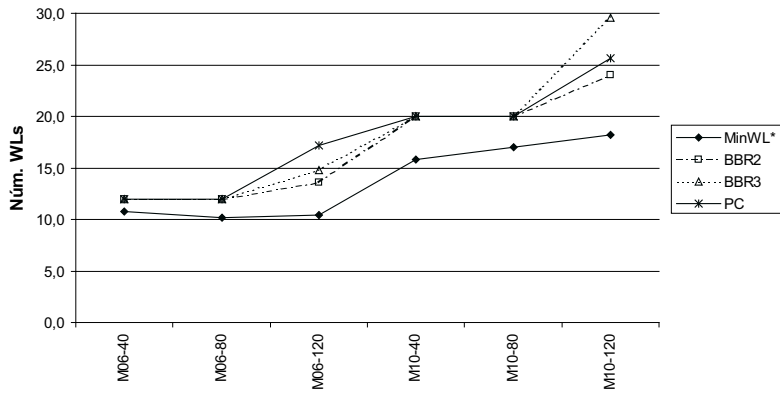
Figura 6.1: Comparação dos Resultados Médios do Método de *Planos-de-Corte* para Redes em Anel



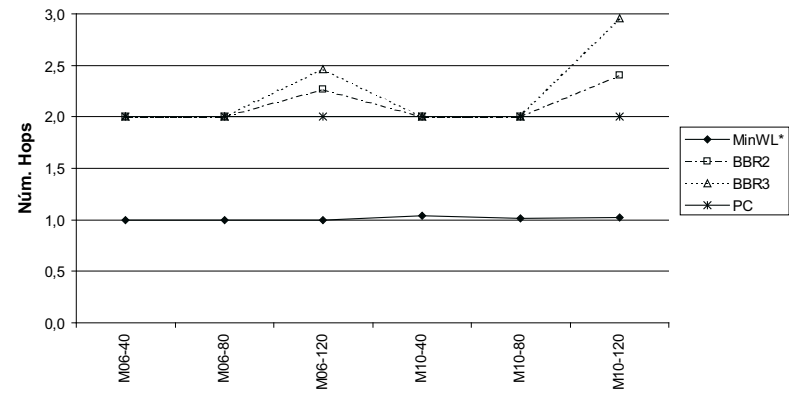
(a) Comparação do Número de ADMs



(b) Comparação do Número de LPs



(c) Comparação do Número de WLS



(d) Comparação do Número de Hops

Figura 6.2: Comparação dos Resultados Médios do Método de *Planos-de-Corte* para Redes Irregulares

Tabela 6.7: Resultados do Algoritmo de *Branch-and-Cut* para Redes em Anel (valores médios)

Instância	Iter	Tempo (seg)	Nós	Total <i>Cuts</i>	<i>Tight</i> (%)	ADMs	LPs	WLs
C04-40	86	1,73	1	332,2	23,5	4,0	4,0	8,0
C04-80	222	4,52	1	532,0	22,1	4,0	4,0	8,0
C04-120	6451	54,54	18	698,8	31,2	4,4	4,2	9,2
C05-40	27	1,55	0	1105,6	7,0	5,0	5,0	10,0
C05-80	29	6,44	0	1807,8	9,0	5,0	5,0	10,0
C05-120	34738	265,83	93	744,6	30,5	5,4	5,2	11,6
C06-40	109	3,82	1	402,0	20,6	6,0	6,0	12,0
C06-80	171	14,74	1	2334,8	9,8	6,0	6,0	12,0
C06-120	40348	400,66	87	1751,4	15,8	6,0	6,0	16,8
C09-40	87	12,93	1	3721,8	4,7	9,0	9,0	18,0
C09-80	40	24,46	1	10154,8	2,8	9,0	9,0	18,0
C09-120	135696	1818,46	221	3032,6	14,3	9,8	9,4	29,6
C10-40	254	19,79	1	884,6	21,3	10,0	10,0	20,0
C10-80	182	49,45	1	13710,2	2,5	10,0	10,0	20,0
C10-120	70639	1157,78	104	9117,4	4,2	10,4	10,2	28,0

6.6.2 Resultados do Algoritmo de *Branch-and-Cut*

O algoritmo de *branch-and-cut* descrito anteriormente também foi implementado em C++, utilizando a biblioteca de funções do COIN-CBC. Os experimentos computacionais também foram realizados na mesma plataforma utilizada anteriormente.

Novamente, os experimentos foram realizados com as mesmas instâncias descritas no capítulo anterior (150 problemas divididos em 3 conjuntos), sendo que estabeleceu-se um limite máximo de tempo de 1800 segundos. Além disso, de forma análoga ao ocorrido durante os testes com o método de *planos-de-corte*, não conseguiu obter resultados satisfatórios quando o algoritmo de *branch-and-cut* foi aplicado às grandes instâncias do problema devido ao número excessivo de restrições.

A Tabela 6.7 apresenta os resultados obtidos pelo algoritmo de *branch-and-cut* para as redes em anel. A primeira coluna identifica o subconjunto de instâncias. Os valores médios para o número de iterações, para o tempo gasto (em segundos) e para o número de nós da árvore (sem computar a raiz) são exibidos na segunda, terceira e quarta colunas, respectivamente. As duas colunas seguintes exibem o número médio de cortes gerados e o percentual de cortes “justos” (ou *tight*) ao final da execução do algoritmo. Por fim, as três últimas colunas exibem os valores médios para o número de multiplexadores *add-drop*, para o número de *lightpaths* e para o número de *wavelength-links* das melhores soluções obtidas pelo algoritmo de *branch-and-cut*.

Observa-se que, apesar da geração de um elevado número de cortes (semelhante ao que ocorreu durante a aplicação do método de *planos-de-corte*), apenas uma pequena fração deles está “ativa” (*tight*) na solução final. Contudo, de modo a se analisar mais detalhadamente esse fato, foi necessário a realização de testes adicionais (descritos na próxima subseção).

Tabela 6.8: Resultados do Algoritmo de *Branch-and-Cut* para Redes Irregulares (valores médios)

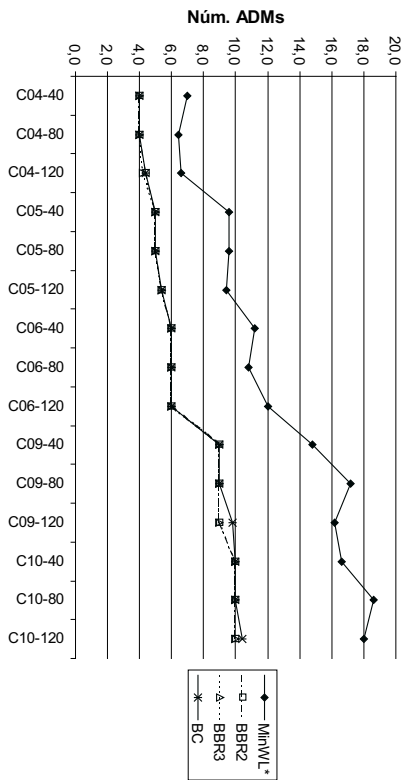
Instância	Iter	Tempo (seg)	Nós	Total <i>Cuts</i>	<i>Tight</i> (%)	ADMs	LPs	WLs
M06-40	369	23,38	1	1100,6	16,6	6,0	6,0	12,0
M06-80	955	78,66	1	1257,8	22,9	6,0	6,0	12,0
M06-120	91660	1222,61	87	1420,4	19,1	6,2	6,2	14,8
M10-40	651	55,64	1	6908	4,0	10,0	10,0	20,0
M10-80	1015	222,94	1	6167,6	7,6	10,0	10,0	20,0
M10-120	45083	1306,63	34	5805,4	9,3	10,4	10,2	25,2

Apesar disso, o algoritmo de *branch-and-cut* foi capaz de encontrar (dentro do limite de tempo estipulado) a solução ótima para 67 das 75 instâncias das redes em anel (89%). Não se obteve a solução ótima apenas para 5 instâncias da rede em anel com 9 nós e 120 produtos e 3 instâncias da rede em anel com 10 nós e 120 produtos.

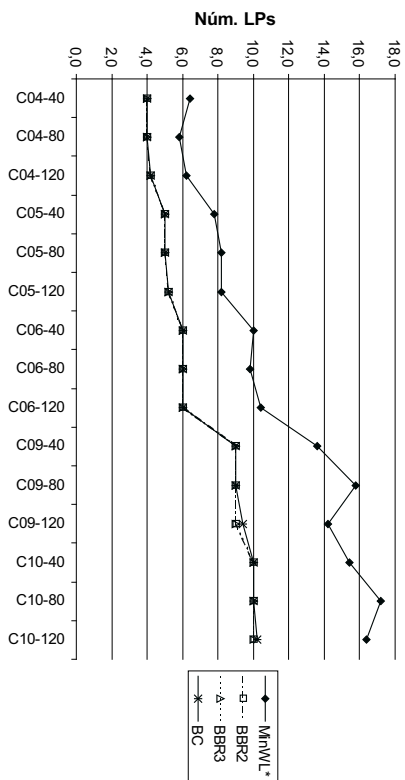
Já a Tabela 6.8 apresenta os resultados obtidos pelo algoritmo de *branch-and-cut* para as redes irregulares. As colunas nessa tabela são as mesmas da tabela anterior. Observa-se que, apesar de um ligeiro aumento dos valores médios de tempo gasto, a proporção de cortes “ativos” na solução final se torna cada vez menor a medida que o tamanho da topologia de rede aumenta (apesar do elevado número de cortes gerados). Mesmo assim, para as redes irregulares, o algoritmo de *branch-and-cut* foi capaz de encontrar (dentro do limite de tempo estipulado) a solução ótima para 25 das 30 instâncias testadas (83%). No total, o algoritmo de *branch-and-cut* obteve a solução ótima para 92 das 105 instâncias (87%), o que impacta diretamente na qualidade das soluções obtidas (se comparadas com aquelas produzidas por outras abordagens).

Os gráficos das Figuras 6.3 e 6.4 apresentam uma comparação dos valores médios obtidos pelo algoritmo de *branch-and-cut* com aqueles obtidos (e apresentados no capítulo anterior) pela heurística proposta em [166] – utilizando a política **MinWL***, bem como com os resultados dos algoritmos de *branch-and-bound* **BBR2** e **BBR3** para as redes em anel e para as redes irregulares, respectivamente. Nesses gráficos, os resultados obtidos pelo algoritmo de *branch-and-cut* estão indicados na legenda pela sigla **BC**.

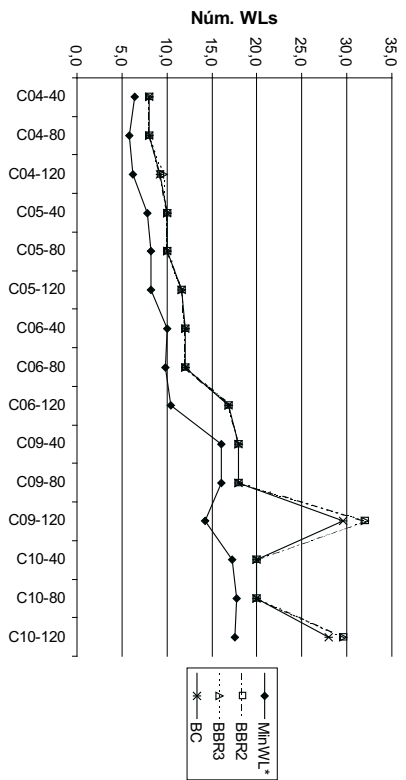
Observa-se que os resultados obtidos pelo algoritmo de *branch-and-cut* são praticamente idênticos aos obtidos pelos algoritmos de *branch-and-bound* **BBR2** e **BBR3** (o que já era esperado em vista de ambas abordagens – poliédrica e lagrangeana – utilizarem a mesma formulação e conseguirem obter soluções ótimas para muitas das instâncias). Contudo, deve-se salientar que **BBR2** obtém seus resultados com um consumo de tempo muito menor (ver Tabela 5.6 e 5.8). Já em relação aos tempos gastos por **BBR3**, verifica-se que eles são mais próximos dos tempos gastos pelo algoritmo de *branch-and-cut* (ver Tabela 5.7 e 5.9). Independentemente disso, ambas as abordagens lagrangeanas são capazes de obter soluções para as grandes instâncias do problema (mesmo que sem garantir sua otimalidade), entretanto a abordagem poliédrica não obteve o mesmo sucesso.



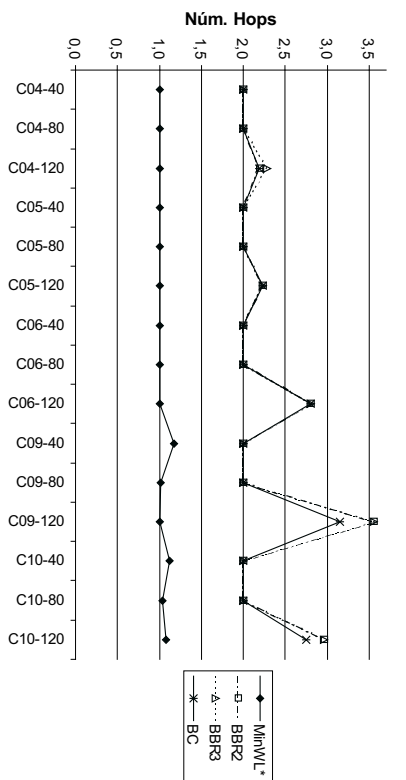
(a) Comparação do Número de ADMs



(b) Comparação do Número de LPs

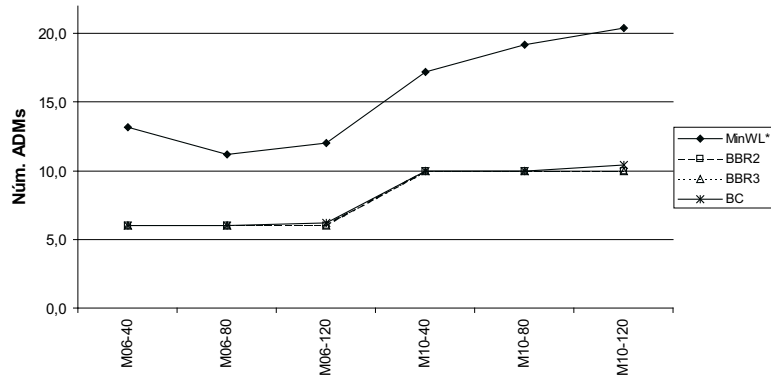


(c) Comparação do Número de WLS

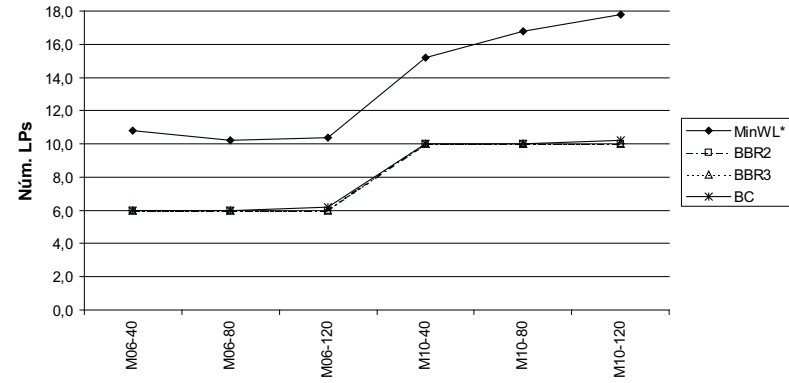


(d) Comparação do Número de Hops

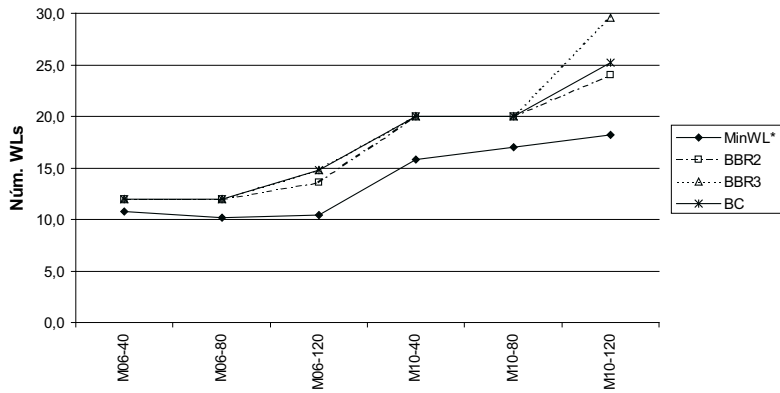
Figura 6.3: Comparação dos Resultados Médios do Algoritmo de *Branch-and-Cut* para Redes em Anel



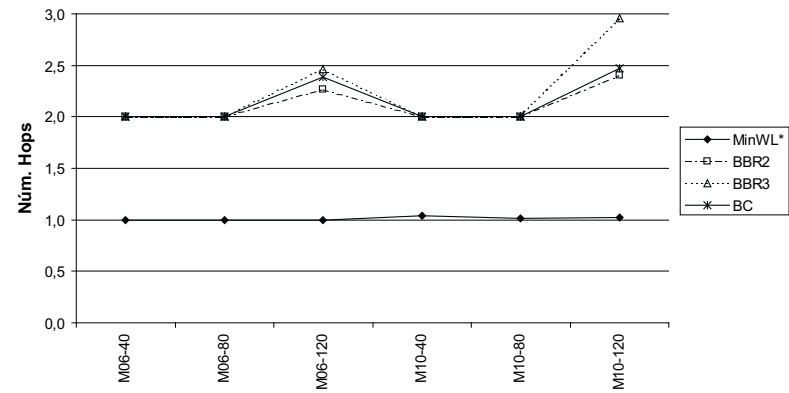
(a) Comparação do Número de ADMs



(b) Comparação do Número de LPs



(c) Comparação do Número de WLS



(d) Comparação do Número de Hops

Figura 6.4: Comparação dos Resultados Médios do Algoritmo de *Branch-and-Cut* para Redes Irregulares

Tabela 6.9: Resultados do Algoritmo de *Cut-and-Branch* para Redes em Anel (valores médios)

Instância	Iter	Tempo (seg)	Nós	Total <i>Cuts</i>	<i>Tight</i> (%)	ADMs	LPs	WLs
C04-40	89	1,98	1	332	23,0	4,0	4,0	8,0
C04-80	169	4,87	1	532	23,9	4,0	4,0	8,0
C04-120	5769	58,32	14	697	31,9	4,4	4,2	9,2
C05-40	27	1,70	0	1106	18,7	5,0	5,0	10,0
C05-80	32	7,55	0	1808	19,8	5,0	5,0	10,0
C05-120	32137	310,05	73	738	32,1	5,4	5,2	11,6
C06-40	129	5,01	1	402	20,5	6,0	6,0	12,0
C06-80	132	16,90	1	2335	20,6	6,0	6,0	12,0
C06-120	51341	573,85	106	1745	19,8	6,0	6,0	16,8
C09-40	112	14,10	1	3722	16,7	9,0	9,0	18,0
C09-80	52	25,29	1	10155	11,2	9,0	9,0	18,0
C09-120	120199	1817,15	155	2302	23,4	10,4	9,8	27,2
C10-40	297	23,22	1	885	22,3	10,0	10,0	20,0
C10-80	160	53,89	1	13710	8,5	10,0	10,0	20,0
C10-120	64196	1165,94	66	5250	16,8	10,4	10,2	27,6

6.6.3 Resultados do Algoritmo de *Cut-and-Branch*

Conforme descrito anteriormente, o algoritmo de *cut-and-branch* pode ser visto como uma implementação restrita de um algoritmo de *branch-and-cut* em que se utilizou o procedimento de geração de novos cortes apenas durante a exploração do nó inicial (ou raiz). Ele também foi implementado em C++, utilizando a biblioteca de funções do COIN-CBC e os experimentos computacionais também foram realizados na mesma plataforma utilizada anteriormente. A motivação para realização desses experimentos foi tentar avaliar os reais “ganhos” obtidos pela utilização do procedimento de geração de novos cortes durante a exploração dos demais nós da árvore.

Novamente, os experimentos foram realizados com as mesmas instâncias descritas no capítulo anterior (150 problemas divididos em 3 conjuntos) e o mesmo limite máximo de tempo (1800 segundos). Além disso, de forma análoga ao ocorrido durante os testes com o algoritmo de *branch-and-cut*, não conseguiu obter resultados satisfatórios quando o algoritmo de *cut-and-branch* foi aplicado às grandes instâncias do problema devido ao número excessivo de variáveis e restrições.

As Tabelas 6.9 e 6.10 apresentam os resultados obtidos pelo algoritmo de *cut-and-branch* para as redes em anel e para as redes irregulares, respectivamente. As colunas nessas tabelas são as mesmas das tabelas anteriores com os resultados do algoritmo de *branch-and-cut* (ver Tabelas 6.7 e 6.8).

Observa-se que, apesar da maioria dos valores apresentados nessas duas tabelas sofrerem poucas alterações em relação àqueles apresentados para o algoritmo de *branch-and-cut*, o percentual de cortes “ativo” aumentou consideravelmente se comparado aos valores obtidos pelo algoritmo de *branch-and-cut*. Uma comparação mais cuidadosa do número de cortes gerados parece indicar que o procedimento de geração

Tabela 6.10: Resultados do Algoritmo de *Cut-and-Branch* para Redes Irregulares (valores médios)

Instância	Iter	Tempo (seg)	Nós	Total <i>Cuts</i>	<i>Tight</i> (%)	ADMs	LPs	WLs
M06-40	387	26,21	1	1101	25,2	6,0	6,0	12,0
M06-80	966	84,55	1	1258	28,5	6,0	6,0	12,0
M06-120	94490	1475,70	57	1176	25,9	6,6	6,4	18,4
M10-40	773	65,26	1	6908	15,7	10,0	10,0	20,0
M10-80	1202	258,65	1	6168	13,1	10,0	10,0	20,0
M10-120	41381	1255,15	25	4845	18,1	10,6	10,4	22,8

de novos cortes (quando aplicado durante o processo de exploração dos demais nós da árvore) conseguiu obter apenas um pequeno conjunto de novos planos de corte (sem contudo produzir melhorias significativas no tempo gasto ou na qualidade das soluções obtidas).

Já os gráficos das Figuras 6.5 e 6.6 apresentam uma comparação dos valores médios obtidos pelo algoritmo de *cut-and-branch* com aqueles obtidos (e apresentados no capítulo anterior) pela heurística proposta em [166] – utilizando a política **MinWL***, bem como com os resultados dos algoritmos de *branch-and-bound* **BBR2** e **BBR3** para as redes em anel e para as redes irregulares, respectivamente. Nesses gráficos, os resultados obtidos pelo algoritmo de *cut-and-branch* estão indicados na legenda pela sigla **CutB**.

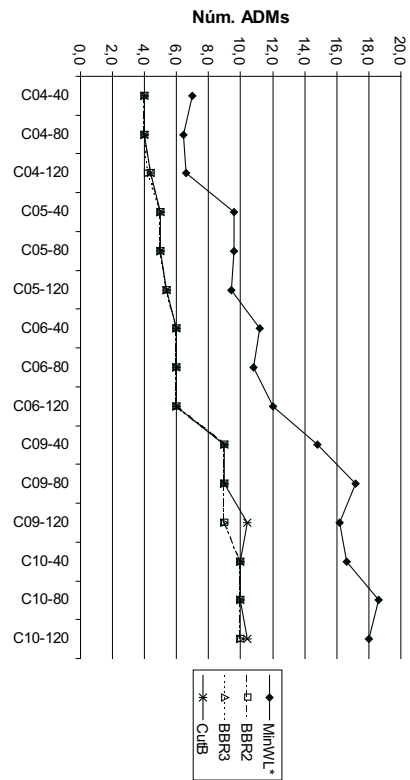
Percebe-se facilmente, pelos gráficos das Figuras 6.5 e 6.6, que a qualidade das soluções obtidas praticamente não foi afetada. Em verdade, o algoritmo de *cut-and-branch* foi capaz de obter a solução ótima para 90 das 105 instâncias testadas (85%), ao passo que o algoritmo de *branch-and-cut* conseguiu encontrar a solução ótima para 92 das instâncias (87%). Isso leva a concluir que o uso do procedimento de geração de novos cortes durante o processo de exploração dos demais nós da árvore é pouco eficaz (pelo menos, nas situações em que os tipos de cortes gerados sejam os mesmos descritos anteriormente).

6.7 Considerações Finais

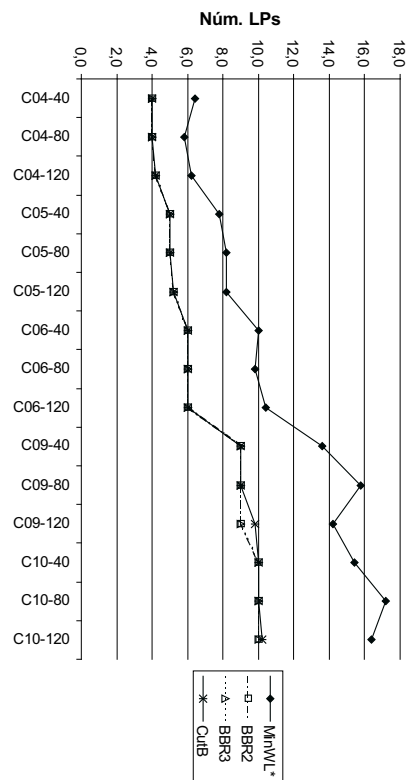
Neste capítulo, procurou-se apresentar os resultados obtidos a partir de uma abordagem poliédrica para resolução do **TGP**. Como os testes iniciais utilizando cortes genéricos se mostraram ineficazes, optou-se pela geração de *planos-de-corte* através da investigação da estrutura facial do poliedro associado ao conjunto de soluções do problema. Os resultados desse estudo foram descritos neste capítulo juntamente com uma avaliação preliminar da “qualidade” dos cortes gerados.

Apresentou-se também neste capítulo os detalhes pertinentes à implementação de um método de *planos-de-cortes* e de um algoritmo de *branch-and-cut* para resolução do **TGP**.

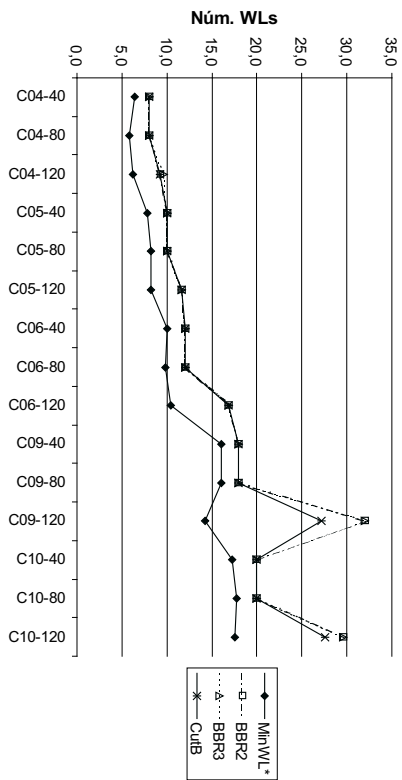
Os resultados obtidos pelo método de *planos-de-cortes* para cerca de uma centena de instâncias foram apresentados e analisados. Esses resultados confirmaram a “boa” qualidade dos cortes utilizados, bem como a “boa” qualidade das soluções obtidas



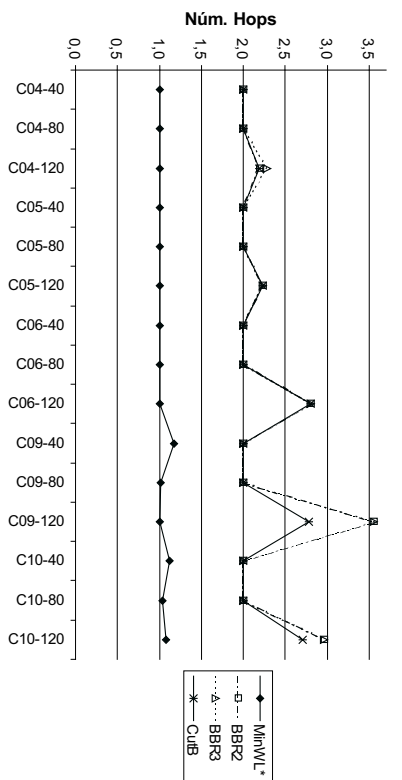
(a) Comparação do Número de ADMs



(b) Comparação do Número de LPs

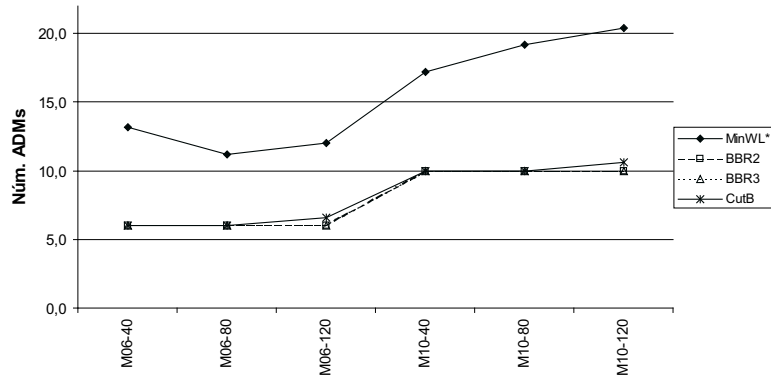


(c) Comparação do Número de WLS

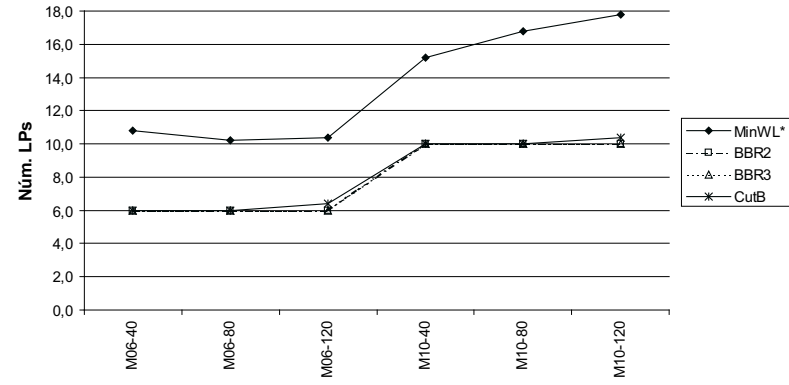


(d) Comparação do Número de Hops

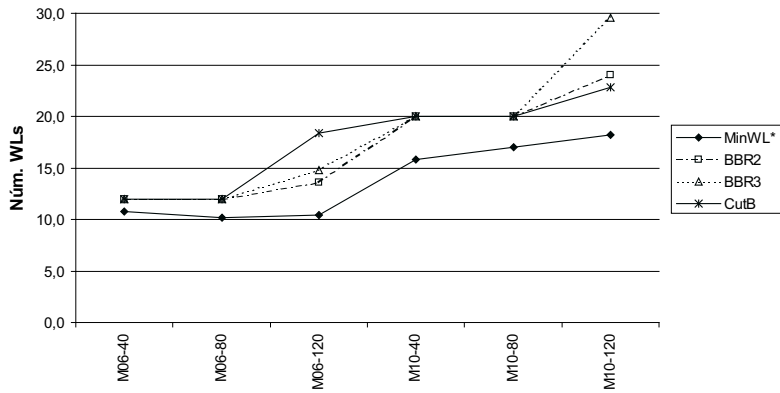
Figura 6.5: Comparação dos Resultados Médios do Algoritmo de *Cut-and-Branch* para Redes em Anel



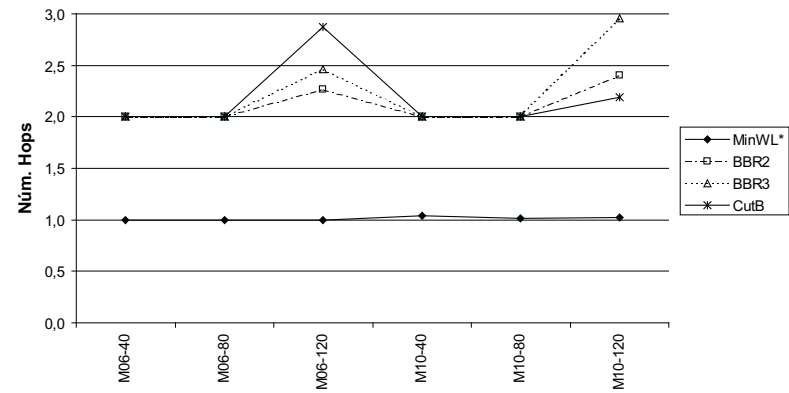
(a) Comparação do Número de ADMs



(b) Comparação do Número de LPs



(c) Comparação do Número de WLS



(d) Comparação do Número de Hops

Figura 6.6: Comparação dos Resultados Médios do Algoritmo de *Cut-and-Branch* para Redes Irregulares

(quando comparadas àquelas produzidas por outras abordagens).

Apresentou-se também os resultados obtidos pelo algoritmo de *branch-and-cut* para cerca de uma centena de instâncias. Os resultados obtidos apontam, pelo menos inicialmente, para uma superioridade da abordagem lagrangeana (do algoritmo de branch-and-bound **BBR2**, em particular), tendo em vista que ambas as abordagens – poliédrica e lagrangeana – obtêm soluções de qualidade equivalente, entretanto o algoritmo **BBR2** apresenta um desempenho muito melhor (isto é, consumo de tempo menor).

Deve-se, ainda, mencionar o fato das abordagens poliédricas não conseguirem lidar adequadamente com as grandes instâncias do problema, em virtude do elevado número de variáveis e restrições presentes quer na formulação original quer nas reformulações geradas a partir da adição dos cortes.

A geração de limites superiores representa um outro ponto a ser enfocado com especial atenção. Experimentos iniciais realizados com heurísticas simples de arredondamento e buscas locais “triviais” para obtenção de soluções viáveis se mostraram incapazes de produzir resultados satisfatórios mesmo para as instâncias pequenas com um alto volume de tráfego (por exemplo, para as redes em anel e irregular com 6 nós). Devido à estrutura especial do **TGP**, optou-se pelo uso de uma heurística combinatória eficiente semelhante àquela utilizada em conjunto com as abordagens lagrangeanas do capítulo anterior. Contudo, a inicialização dessa heurística é realizada através do arredondamento da solução corrente da relaxação linear (diferentemente das abordagens lagrangeanas, em que a heurística é inicializada com a solução obtida para o limite inferior que, por sua vez, já é inteira). Esse procedimento, infelizmente, não parece ser tão eficaz, uma vez que para alguns subconjuntos de instâncias (por exemplo, a rede em anel com 10 nós e 120 produtos), apesar do limite inferior ao término da busca já estar bem próximo da solução ótima do problema, o algoritmo de *branch-and-cut* não conseguiu provar a otimalidade das soluções obtidas para tais instâncias.

Por fim, apresentou-se também neste capítulo os resultados obtidos pelo algoritmo de *cut-and-branch* para cerca de uma centena de instâncias. Esse algoritmo pode ser visto como uma implementação restrita do algoritmo de *branch-and-cut* em que o procedimento de geração de cortes é aplicado apenas durante a exploração do nó inicial. Os resultados obtidos indicam que o uso do procedimento de geração de novos cortes durante o processo de exploração dos demais nós da árvore é pouco eficaz (pelo menos, nas situações em que os tipos de cortes gerados sejam os mesmos descritos anteriormente). Mais ainda, os experimentos demonstraram que a qualidade das soluções obtidas pelo algoritmo de *cut-and-branch* era praticamente idêntica a das soluções obtidas pelo algoritmo de *branch-and-cut*. Contudo, deve-se mencionar que um processo de busca mais ágil (por exemplo, sem a utilização do procedimento de geração de cortes a cada nó explorado) pode vir a viabilizar a resolução de instâncias maiores no futuro (desde que se contorne as dificuldades relacionadas ao elevado número de variáveis e restrições).

Capítulo 7

Comparação entre Abordagens do TGP

NESTE capítulo, apresenta-se uma análise comparativa entre as duas abordagens propostas para resolução do **TGP** – lagrangeana e poliédrica. De modo a viabilizar os experimentos com ambas as abordagens em igualdade de condições, um novo conjunto de instâncias para teste foi gerado. A seção 7.1 fornece detalhes sobre a construção de tais instâncias. Os resultados obtidos pelos experimentos computacionais utilizando-se as diversas abordagens propostas para resolução do **TGP** são apresentados na seção 7.2. Por fim, a seção 7.3 apresenta algumas considerações finais.

7.1 Descrição dos Experimentos

Um conjunto de instâncias para realização de novos testes envolvendo todas as abordagens descritas anteriormente foi construído de forma semelhante às instâncias utilizadas em [166].

Tendo em vista dificuldades enfrentadas pela abordagem poliédrica para resolução de grandes instâncias, de modo a se procurar (ou, pelo menos, tentar) garantir igualdade de condições entre as abordagens, utilizou-se na construção dessas instâncias a rede irregular de 6 nós cuja topologia física é apresentada na Figura 5.3(a), semelhante ao realizado em [166].

Ainda seguindo as orientações de [166], cinco matrizes com 50 requisições de tráfego não-uniforme foram geradas aleatoriamente da seguinte forma. Considerou-se a utilização de três tipos distintos de conexões: 51,84 Mbps (OC-1), 155,52 Mbps (OC-3) e 622,08 Mbps (OC-12). Mais ainda, a quantidade (em canais) de tráfego de cada requisição para um dado tipo de conexão foi distribuída uniformemente nos seguintes intervalos:

- Conexões de 51,84 Mbps (OC-1): entre 0 e 16 canais por cada requisição de tráfego;
- Conexões de 155,52 Mbps (OC-3): entre 0 e 8 canais por cada requisição de tráfego;

- Conexões de 622,08 Mbps (OC-12): entre 0 e 2 canais por cada requisição de tráfego.

O volume total de tráfego de cada matriz era em média equivalente a 628 requisições de 51,84 Mbps (OC-1). Além disso, a capacidade de transporte de cada comprimento de onda foi feita equivalente a 2488,32 Mbps (OC-48) e nos testes realizados foram necessários 4 comprimentos de onda distintos para se obter, pelo menos, uma solução viável (o que equivale a uma capacidade total capaz de transportar 192 conexões de 51,84 Mbps para cada fibra e sentido).

7.2 Avaliação dos Resultados

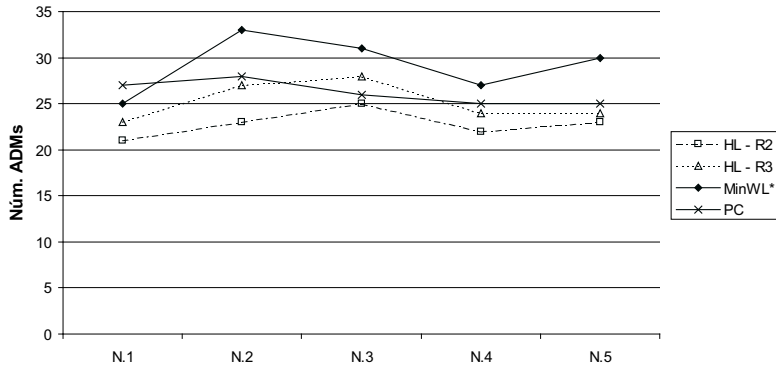
Os gráficos da Figura 7.1 exibem os resultados obtidos pelas diversas abordagens heurísticas apresentadas nos capítulos anteriores para cada uma das cinco instâncias descritas anteriormente. Nesses gráficos, os resultados obtidos pela heurística lagrangeana estão indicados na legenda pelas siglas **HL-R2** e **HL-R3** de acordo com a relaxação utilizada (relaxação #2 ou #3). Já os resultados obtidos pelo método de *planos-de-corte* estão indicados na legenda pela sigla **PC**, enquanto que os resultados da heurística proposta em [166], pela sigla **MinWL***.

Observa-se que a heurística lagrangeana utilizando a relaxação #2 apresenta melhores resultados para o número de multiplexadores de *add-drop* e de *lightpaths*. Já os resultados obtidos pela heurística lagrangeana utilizando a relaxação #3 e pelo método de *planos-de-corte* são muito próximos (alternando de acordo com a instância). Além disso, na maioria dos casos, eles não são muito distantes daqueles obtidos pela heurística lagrangeana utilizando a relaxação #2. Finalmente, a heurística proposta em [166] apresenta os piores resultados (com exceção da primeira instância, em que ela supera apenas o método de *planos-de-corte*).

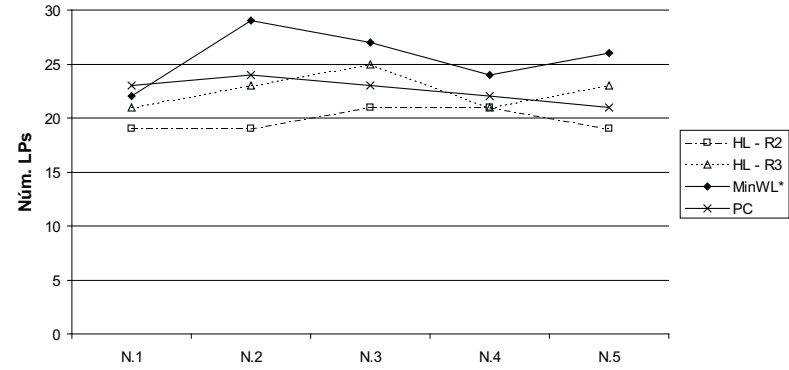
Examinando-se os resultados relativos aos números de *wavelength-links* e de “saltos” (hops), verifica-se que, semelhante ao ocorrido anteriormente, a heurística proposta em [166] optou por utilizar um número menor de “saltos” na topologia virtual e, consequentemente, necessitou de uma quantidade maior de multiplexadores de *add-drop* e de *lightpaths* (para a segunda instância esses valores chegam a ser 40% acima dos necessários à heurística lagrangeana utilizando a relaxação #2).

No tocante as demais abordagens, observa-se novamente o que já havia sido constatado em testes anteriores. A melhor utilização dos recursos de rede (obtendo soluções com um número menor de ADMs e *lightpaths*) está associada a um número maior de “saltos” (*hops*) na topologia virtual para rotear as requisições de tráfego.

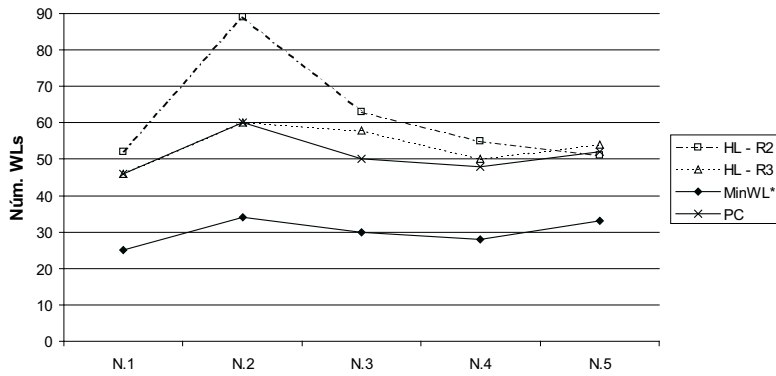
Contudo, esse problema poderia ser contornado de duas formas distintas. Primeiramente, a formulação apresentada para o TGP poderia ser alterada, de modo que os tamanhos dos *lightpaths* fossem contabilizados e, eventualmente, limitados. Entretanto, essa abordagem implicaria em uma expansão da formulação, inviabilizando, eventualmente, a sua resolução. Uma segunda forma de lidar com o problema mencionado seria manipular o valor dos custos associados aos arcos entre elementos distintos da rede, procurando-se, assim, desestimular o estabelecimento de “longos” *lightpaths*. Todavia, esse artifício, apesar de aparentemente simples, pode se mostrar de difícil realização (uma vez que não se sabe ao certo como estabelecer tais custos),



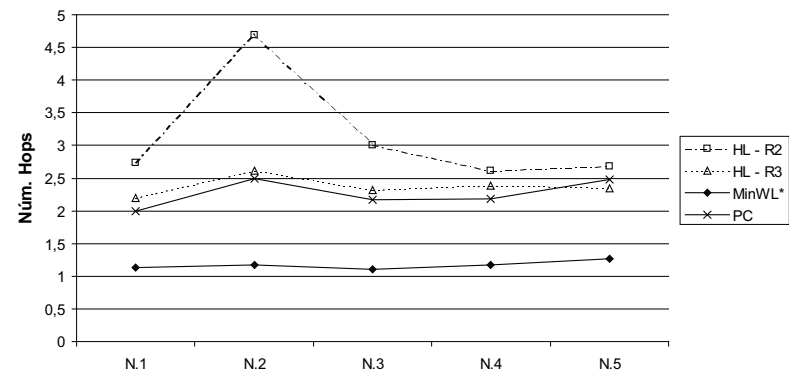
(a) Comparação do Número de ADMs



(b) Comparação do Número de LPs



(c) Comparação do Número de WLs



(d) Comparação do Número de Hops

Figura 7.1: Comparação dos Resultados entre Abordagens Heurísticas

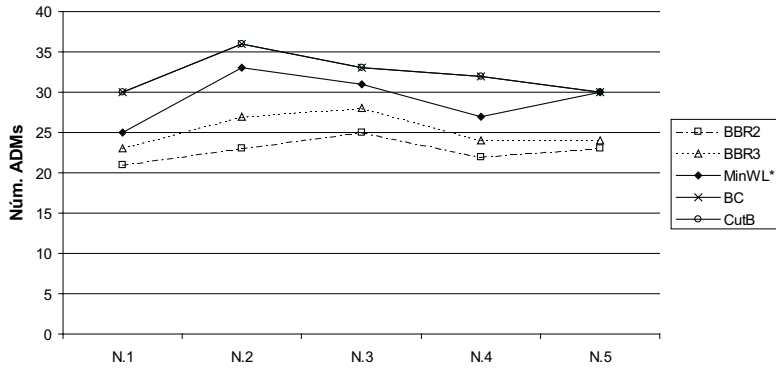
além de não garantir efetivamente uma limitação para o tamanho dos *lightpaths*.

Já os gráficos da Figura 7.2 exibem os resultados para cada uma das cinco instâncias obtidos pelas diversas abordagens exatas apresentadas anteriormente. Nesses gráficos, os resultados obtidos pelos algoritmos de *branch-and-bound* estão indicados na legenda pelas siglas **BBR2** e **BBR3** de acordo com a relaxação utilizada (relaxação #2 ou #3). Já os resultados obtidos pelos algoritmos de *branch-and-cut* e *cut-and-branch* estão indicados na legenda pelas siglas **BC** e **CutB**, respectivamente. Além disso, apesar da heurística proposta em [166] não representar um método de solução exato para o TGP, seus resultados também foram apresentados nos gráficos da Figura 7.2 apenas para servirem como referência e estão indicados na legenda pela sigla **MinWL***.

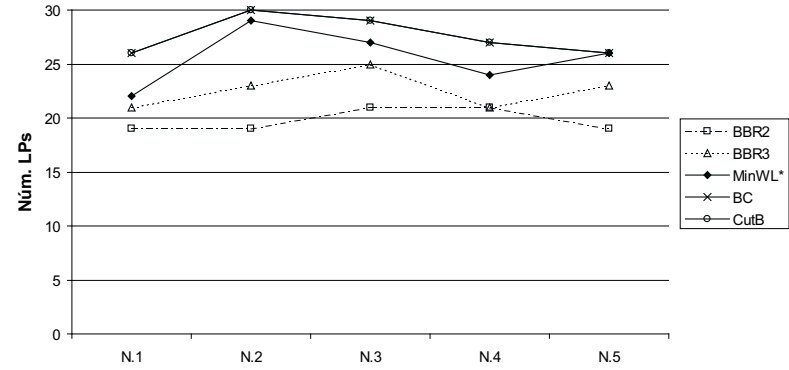
Deve-se destacar que nenhuma das abordagens exatas propostas – poliédricas ou lagrangeanas – conseguiu encontrar soluções ótimas para as instâncias testadas dentro do intervalo de tempo estipulado (1800 segundos). Mesmo assim, as abordagens lagrangeanas obtiveram melhores soluções (sendo que **BBR2** manteve sua superioridade). Já as abordagens poliédricas, por outro lado, não encontraram soluções melhores que a da heurística proposta em [166]. Na realidade, o baixo desempenho dessas abordagens parece estar ligado diretamente às dificuldades em se obter “boas” soluções viáveis (conforme discutido ao final do capítulo 6) a partir da relaxação linear do problema. Infelizmente, apesar dos esforços realizados, não se conseguiu obter nenhuma melhoria no sentido de se produzir soluções viáveis de qualidade superior àquelas já encontradas.

No tocante as demais abordagens, observa-se novamente o que já havia sido constatado em testes anteriores. Deve-se destacar que, semelhante ao ocorrido no capítulo 5, a heurística lagrangeana conseguiu obter praticamente as mesmas soluções que os algoritmos de *branch-and-bound*. Contudo, estes últimos consumiram todo o limite de tempo dado sem conseguir provar a otimalidade das soluções encontradas. O algoritmo de *branch-and-bound* **BBR2** encerrou a busca (após explorar alguns milhares de nós) com um gap de dualidade acima de 20% para todas as instâncias. Já no algoritmo de *branch-and-bound* **BBR3**, quando a busca foi encerrada o *gap de dualidade* para todas as instâncias estava acima de 30%. Já a heurística lagrangeana obteve os resultados apresentados nos gráficos da Figura 7.1 em menos de 50 segundos.

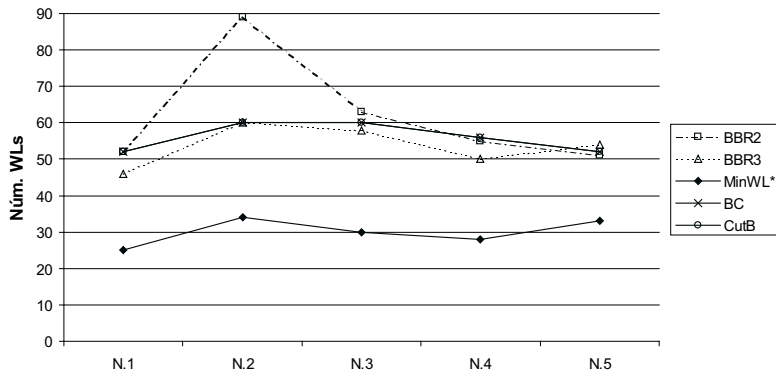
Finalmente, os resultados indicam que as abordagens testadas enfrentam maiores dificuldades quando o número de tipos de conexão aumenta. No capítulo 5, obteve-se soluções ótimas (ou “quase-ótimas”) para várias instâncias cujas topologias possuem um número de nós de 2,5 a 3 vezes maior que a rede descrita na seção 7.1 e para um volume de tráfego de 120 requisições. Isto se deve, pelo menos em parte, ao fato de se utilizar naqueles testes apenas um tipo de conexão, contribuindo assim para a redução do número de combinações possíveis e facilitando a obtenção de uma solução. Na medida que o número de tipos de conexão cresce, aumenta também o número de possíveis combinações para obtenção de uma solução – dificultando, conseqüentemente, a prova de sua otimalidade. Nestes casos, a heurística lagrangeana parece ser mais indicada, uma vez que ela se mostrou capaz de obter bons resultados numa fração do tempo gasto pelas abordagens exatas.



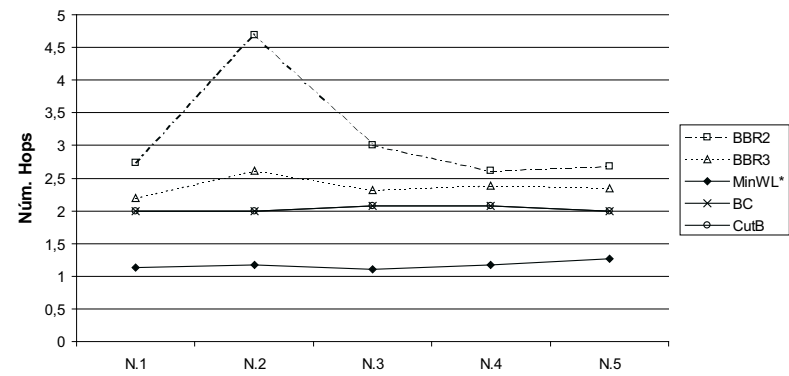
(a) Comparação do Número de ADMs



(b) Comparação do Número de LPs



(c) Comparação do Número de WLs



(d) Comparação do Número de Hops

Figura 7.2: Comparação dos Resultados entre Abordagens Exatas

7.3 Considerações Finais

Neste capítulo, apresentou-se os resultados obtidos durante os testes de todas as abordagens propostas neste trabalho utilizando um conjunto de instâncias razoavelmente difíceis.

Além de validar as conclusões apresentadas anteriormente, o maior objetivo desses experimentos foi *reapresentar* de forma sucinta uma série de conclusões já obtidas anteriormente e apresentadas de forma dispersa nos capítulos 5 e 6.

Capítulo 8

Variações do TGP

8.1 Introdução

NESTE capítulo é apresentada uma abordagem para resolução do **TGRP** e do **STGP** através da relaxação lagrangeana, mais especificamente, através do uso de uma heurística lagrangeana, em vista dos “bons” resultados apresentados anteriormente. A seção 8.2 apresenta um abordagem lagrangeana para resolução do **TGRP**, enquanto que o **STGP** é explorado na seção 8.3. Por fim a seção 8.4 apresenta algumas considerações finais.

8.2 Abordagem Lagrangeana do TGRP

Nesta seção é apresentada uma abordagem para resolução do **TGRP** através da relaxação lagrangeana. Na próxima subseção a relaxação lagrangeana é utilizada para obtenção de um limite inferior para a formulação do problema $TGRP_C$ apresentada na subseção 3.4.2. Tal limite é utilizado, juntamente com uma heurística para geração de limites superiores descrita na seção 8.2.2, na implementação de uma heurística lagrangeana para a resolução do problema (ver subseções 8.2.3 e 8.2.4). Por fim, a subseção 8.2.5 apresenta os resultados obtidos nos experimentos computacionais (apresentados também em [122, 147]).

8.2.1 Limite Inferior para TGRP

De modo a se realizar a relaxação lagrangeana da formulação proposta, um vetor de multiplicadores de Lagrange λ , tal que $\lambda_\ell^t \in \mathbb{R}_+, \forall \ell \in A_\ell, \forall t \in T$, será associado às restrições que limitam a capacidade dos arcos – restrições (3.19d). Já as restrições (3.19e) serão associadas a um vetor de multiplicadores θ , tal que $\theta_\ell \in \mathbb{R}_+, \forall \ell \in A_\ell$.

Posto isto, será apresentada a seguir a formulação e procedimentos de cálculo de um limite inferior obtidos através do uso da relaxação lagrangeana. Neste limite são relaxadas as restrições (3.19d) que limitam a capacidade dos containers virtuais e que estabelecem uma ligação entre as variáveis de fluxo e as variáveis de decisão, juntamente com as restrições (3.19e) que asseguram que a alocação de um *container* ocorre em um único período.

Relaxação #1

Neste limite, L_1^{TGRP} , são relaxadas as restrições (3.19d) responsáveis por estabelecer uma ligação entre as variáveis de fluxo (f_ℓ^{pt}) e as variáveis de decisão (w_ℓ^t), bem como as restrições (3.19e). Isto resulta na eliminação do relacionamento entre as variáveis de fluxo e de decisão. Mais ainda, as variáveis de fluxo de produtos distintos em períodos distintos também deixam de ser interdependentes uma vez que essas restrições – restrições (3.19d) – eram responsáveis pelo limite sobre o volume total de produtos transportados através de um arco em um determinado período de tempo. Além disso a limitação sobre a alocação dos *containers* – restrições (3.19e) – também são relaxadas, tornando mais fácil o cálculo do limite inferior.

Dessa forma, o limite inferior, L_1^{TGRP} , pode ser formulado da seguinte maneira:

$$L_1^{\text{TGRP}}(\lambda, \theta) = L_{1W}^{\text{TGRP}}(\lambda, \theta) + L_{1F}^{\text{TGRP}}(\lambda) - \sum_{\ell \in A_\ell} \theta_\ell \quad (8.1)$$

em que:

$$L_{1W}^{\text{TGRP}}(\lambda, \theta) = \min \sum_{t \in T} \sum_{\ell \in A_\ell} (c_\ell^t + \theta_\ell - \Delta \sum_{t'=t}^{n_t} \lambda_\ell^{t'}) w_\ell^t \quad L_{1F}^{\text{TGRP}}(\lambda) = \min \sum_{t \in T} \sum_{t_p \in P} \sum_{\ell \in A_\ell} \delta_p \lambda_\ell^t f_\ell^{pt}$$

sujeito a: (3.19c) e (3.19g).

sujeito a: (3.19b), (3.19f) e (3.19h).

O subproblema $L_{1F}^{\text{TGRP}}(\lambda)$ pode ser decomposto em $n_p n_t = |P| \cdot |T|$ problemas independentes de modo que:

$$L_{1F}^{\text{TGRP}}(\lambda) = \sum_{t \in T} \sum_{t_p \in P} \delta_p L_{1F_{pt}}^{\text{TGRP}}(\lambda) \quad (8.2)$$

em que para cada produto $t_p \in P$ e para cada período $t \in T$:

$$L_{1F_{pt}}^{\text{TGRP}}(\lambda) = \min \sum_{\ell \in A_\ell} \lambda_\ell^t f_\ell^{pt} \quad (8.3a)$$

sujeito a:

$$\sum_{\ell \in A_\ell^+(i)} f_\ell^{pt} - \sum_{\ell \in A_\ell^-(i)} f_\ell^{pt} = b_i^{pt}, \quad \forall i \in N_\ell \quad (8.3b)$$

$$f_\ell^{pt} \geq 0, \quad \forall \ell \in A_\ell \quad (8.3c)$$

$$f_\ell^{pt} \text{ inteiro}, \quad \forall \ell \in A_\ell \quad (8.3d)$$

As restrições de integralidade (8.3d) podem ser desprezadas, uma vez que a matriz de coeficientes do conjunto restante de restrições, (8.3b)–(8.3c), é totalmente unimodular. Além disso, cada um dos subproblemas $L_{1F_{pt}}^{\text{TGRP}}(\lambda)$ pode ser resolvido através do cálculo do caminho mínimo entre o nó de origem s_p e o nó de destino d_p de cada produto.

Seja \mathcal{C}^{pt} o conjunto de arcos pertencentes ao caminho mínimo entre s_p e d_p calculado utilizando-se λ_ℓ^t como custo de cada arco. Sendo assim, a solução do segundo subproblema pode ser obtida roteando-se o tráfego de cada produto $t_p \in P$ no período $t \in T$ através do caminho mínimo \mathcal{C}^{pt} .

Considere que $\bar{f}_\ell^{pt}, \forall t_p \in P, \forall \ell \in A_\ell, \forall t \in T$, representam os valores obtidos, dessa maneira, para as variáveis de fluxo. Desse modo o valor da solução do segundo subproblema, $L_{\text{IF}}^{\text{TGRP}}$, será dado por:

$$L_{\text{IF}}^{\text{TGRP}}(\lambda) = \sum_{t \in T} \sum_{t_p \in P} \delta_p \sum_{\ell \in A_\ell} \lambda_\ell^t \bar{f}_\ell^{pt} = \sum_{t \in T} \sum_{t_p \in P} \delta_p f_p^t \left(\sum_{\ell \in C^{pt}} \lambda_\ell^t \right). \quad (8.4)$$

Já o primeiro subproblema, $L_{\text{IW}}^{\text{TGRP}}$, necessita de um procedimento um pouco mais elaborado para sua resolução. Antes, porém, para se obter uma solução relaxada mais próxima da efetiva solução do problema, deve-se agregar as restrições (8.5a) e (8.5b) – inequações válidas – à formulação do subproblema $L_{\text{IW}}^{\text{TGRP}}$ dada anteriormente.

$$\sum_{t'=1}^t \sum_{\ell \in A_\ell^+(i)} w_\ell^{t'} \geq \beta_i^t, \quad \forall i \in N_\ell^e, \forall t \in T, \quad (8.5a)$$

$$\sum_{t'=1}^t \sum_{\ell \in A_\ell^-(i)} w_\ell^{t'} \geq \gamma_i^t, \quad \forall i \in N_\ell^e, \forall t \in T. \quad (8.5b)$$

em que β_i^t e $\gamma_i^t, \forall i \in N_\ell^e, \forall t \in T$, representam o número mínimo de arcos necessários para se transportar, respectivamente, todo o tráfego originado e destinado a cada nó de *add-drop* em um dado período e são dados pelas seguintes expressões:

$$\beta_i^t = \left[\sum_{t_p \in P \mid s_p=i} \frac{f_p^t}{m_p} \right], \quad \forall i \in N_\ell^e, \forall t \in T, \quad \gamma_i^t = \left[\sum_{t_p \in P \mid d_p=i} \frac{f_p^t}{m_p} \right], \quad \forall i \in N_\ell^e, \forall t \in T.$$

Desse modo assegura-se que o número de arcos utilizados na solução obtida seja suficiente para transportar toda a oferta (ou demanda) de um produto originado (ou destinado) a um nó de *add-drop* em um dado período. Em verdade, como antes, pode-se considerar que tais restrições já pertenciam à formulação original do problema – sem contudo modificar sua solução – e que foram mantidas intocadas no modelo relaxado. Considere ainda que as seguintes variáveis de folga serão agregadas às restrições (8.5a) e (8.5b):

s_i^t representa a quantidade de facilidades instaladas nos arcos que saem de $i \in N_\ell^e$ no período $t \in T$ e que será utilizada em um período futuro $t' > t$;

r_i^t representa a quantidade de facilidades instaladas nos arcos que chegam em $i \in N_\ell^e$ no período $t \in T$ e que será utilizada em um período futuro $t' > t$.

Sendo assim, as restrições (8.5a) e (8.5b) podem ser reescritas da seguintes forma:

$$\sum_{t'=1}^t \sum_{\ell \in A_\ell^+(i)} w_\ell^{t'} - s_i^t = \beta_i^t, \quad \forall i \in N_\ell^e, \forall t \in T, \quad (8.6a)$$

$$\sum_{t'=1}^t \sum_{\ell \in A_\ell^-(i)} w_\ell^{t'} - r_i^t = \gamma_i^t, \quad \forall i \in N_\ell^e, \forall t \in T. \quad (8.6b)$$

Subtraindo-se, membro a membro, as equações referentes a dois períodos consecutivos, obtém-se:

$$s_i^{t-1} + \sum_{\ell \in A_\ell^+(i)} w_\ell^t - s_i^t = \bar{\beta}_i^t, \forall i \in N_\ell^e, \forall t \in T, \quad (8.7a)$$

$$r_i^{t-1} + \sum_{\ell \in A_\ell^-(i)} w_\ell^t - r_i^t = \bar{\gamma}_i^t, \forall i \in N_\ell^e, \forall t \in T, \quad (8.7b)$$

em que $s_i^0 = r_i^0 = \beta_i^0 = \gamma_i^0 = 0, \forall i \in N_\ell^e$, enquanto que: $\bar{\beta}_i^t = \beta_i^t - \beta_i^{t-1}, \forall i \in N_\ell^e, \forall t \in T$ e $\bar{\gamma}_i^t = \gamma_i^t - \gamma_i^{t-1}, \forall i \in N_\ell^e, \forall t \in T$.

Portanto, L_{1W}^{TGRP} pode ser reescrito da seguinte maneira:

$$L_{1W}^{\text{TGRP}}(\lambda, \theta) = \min \sum_{t \in T} \sum_{\ell \in A_\ell} (c_\ell^t + \theta_\ell - \Delta \sum_{t'=t}^{n_t} \lambda_\ell^{t'}) w_\ell^t$$

sujeito a:

$$\sum_{\ell \in A_\ell^+(i)} w_\ell^t - \sum_{\ell \in A_\ell^-(i)} w_\ell^t = 0, \forall i \in N_\ell^o, \forall t \in T \quad (8.8b)$$

$$s_i^{t-1} + \sum_{\ell \in A_\ell^+(i)} w_\ell^t = \bar{\beta}_i^t + s_i^t, \forall i \in N_\ell^e, \forall t \in T \quad (8.8c)$$

$$r_i^{t-1} + \sum_{\ell \in A_\ell^-(i)} w_\ell^t = \bar{\gamma}_i^t + r_i^t, \forall i \in N_\ell^e, \forall t \in T \quad (8.8d)$$

$$w_\ell^t \in \{0, 1\}, \forall \ell \in A_\ell, \forall t \in T \quad (8.8e)$$

$$s_i^t \geq 0, \forall i \in N_\ell^e, \forall t \in T \quad (8.8f)$$

$$r_i^t \geq 0, \forall i \in N_\ell^e, \forall t \in T \quad (8.8g)$$

A formulação anterior torna explícita a natureza dinâmica do primeiro subproblema L_{1W}^{TGRP} . Além disso, pode-se tratá-lo como problema de fluxo de custo mínimo multiperíodo em que existe a possibilidade de armazenar algum estoque (neste caso, um estoque de *containers* representado por s_i^t e r_i^t) para uso no futuro.

Considere que $\bar{w}_\ell^t, \forall \ell \in \tilde{A}_\ell, \forall t \in T$, representam os valores das variáveis de decisão obtidos a partir da resolução do primeiro subproblema. Desse modo o valor da solução do primeiro subproblema, L_{1W}^{TGRP} , será dado por:

$$L_{1W}^{\text{TGRP}}(\lambda, \theta) = \sum_{t \in T} \sum_{\ell \in A_\ell} (c_\ell^t + \theta_\ell - \Delta \sum_{t'=t}^{n_t} \lambda_\ell^{t'}) \bar{w}_\ell^t. \quad (8.9)$$

Portanto, o valor total do primeiro limite inferior, L_1^{TGRP} , será dado por:

$$L_1^{\text{TGRP}}(\lambda, \theta) = \sum_{t \in T} \sum_{\ell \in A_\ell} \left(c_\ell^t + \theta_\ell - \Delta \sum_{t'=t}^{n_t} \lambda_\ell^{t'} \right) \bar{w}_\ell^t + \sum_{t \in T} \sum_{t_p \in P} \delta_p f_p^t \left(\sum_{\ell \in C^{pt}} \lambda_\ell^t \right) - \sum_{\ell \in A_\ell} \theta_\ell. \quad (8.10)$$

8.2.2 Limite Superior para TGRP

A heurística utilizada para cálculo de limites superiores para o **TGRP** é muito semelhante a descrita na seção 5.2 para o **TGP** e pode ser vista, na verdade, como uma extensão daquela heurística.

Seja \hat{A}_0^t o conjunto de facilidades (arcos) disponíveis para o cálculo do limite superior no período $t \in T$. Inicialmente esse conjunto pode ser igual à união dos conjuntos de arcos $\bar{A}_\ell^t = \{\ell \in A_\ell \mid \bar{w}_\ell^t = 1\}$, $\forall t \in T$, selecionados durante o cálculo do limite inferior, isto é, $\hat{A}_0^t = \bigcup_{t'=1}^t \{\ell \in A_\ell \mid \bar{w}_\ell^{t'} = 1\}$, em que $\bar{w}_\ell^{t'}$ representa o valor da variável de decisão associada ao arco $\ell \in A_\ell$ no período $t' \in T$ obtido pelo cálculo do limite inferior, ou, até mesmo, através de uma outra heurística qualquer como, por exemplo, $\hat{A}_0^t = \bigcup_{t'=1}^t \{\ell \in A_\ell \mid \bar{f}_\ell^{pt'} > 0\}$ ou $\hat{A}_0^t = \bigcup_{t'=1}^t \{\ell \in A_\ell \mid \bar{w}_\ell^{t'} = 1 \text{ ou } \bar{f}_\ell^{pt'} > 0\}$.

De forma a se assegurar que as restrições (3.19e) não são violadas na eventualidade de um mesmo arco ter sido utilizado em vários períodos, apenas sua primeira ocorrência será considerada (isto é, somente o período em que foi alocado pela primeira vez será computado na inicialização de \hat{A}_0^t). Entretanto, para os períodos futuros, caso o arco não tenha sido alocado naquele período, deve-se verificar se ele já não foi alocado em um período anterior.

Em seguida, a heurística descrita para o **TGP** (ver seção 5.2) é aplicada para cada período, tomando-se o cuidado de garantir sempre o atendimento das restrições (3.19e) – conforme descrito acima.

8.2.3 Método de Subgradiente para TGRP

Para cada conjunto de multiplicadores de Lagrange, obtém-se um limite inferior para o **TGRP**, dado por $L_1^{\text{TGRP}}(\lambda, \theta)$. De modo a se obter uma solução para o problema, deve-se procurar maximizar o valor do limite inferior obtido.

Neste sentido, semelhante ao que foi descrito na seção 5.3, o melhor limite inferior será dado pela solução do problema lagrangeano dual \bar{P}_1^D , caso a relaxação #1 seja aplicada:

$$(\bar{P}_1^D) \quad \max_{\substack{\theta \geq 0 \\ \lambda \geq 0}} L_1^{\text{TGRP}}(\lambda, \theta). \quad (8.11)$$

Maiores detalhes sobre o método de subgradiente podem ser encontrados na seção 5.3. O cálculo da direção de subida e do tamanho do passo, bem como a próprio algoritmo de subgradiente, necessitam apenas de pequenos ajustes devidos às alterações realizadas para o cálculo do limite inferior de **TGRP** e, portanto, não são apresentados neste texto. A seguir, apresenta-se os subgradientes referentes a relaxação #1.

Na relaxação #1, tem-se um conjunto de multiplicadores (λ, θ) , o qual é utilizado na solução do problema relaxado e, conseqüentemente, na obtenção do limite inferior. Associado à solução relaxada $(\bar{w}_\ell^t, \bar{f}_\ell^{pt})$, tem-se os subgradientes $\xi_1^{\text{TGRP}}(\bar{w}_\ell^t, \bar{f}_\ell^{pt}) =$

$(\xi_{1\ell t}^{\text{TGRP}}(\bar{w}_\ell^t, \bar{f}_\ell^{pt}), \xi_{1\ell}^{\text{TGRP}}(\bar{w}_\ell^t))$ definidos abaixo:

$$\xi_{1\ell t}^{\text{TGRP}}(\bar{w}_\ell^t, \bar{f}_\ell^{pt}) = \sum_{t_p \in P} \delta_p \bar{f}_\ell^{pt} - \Delta \sum_{\tilde{t}=1}^t \bar{w}_\ell^{\tilde{t}}, \forall \ell \in A_\ell, \forall t \in T \quad (8.12a)$$

$$\xi_{1\ell}^{\text{TGRP}}(\bar{w}_\ell^t) = \sum_{t \in T} \bar{w}_\ell^t - 1, \forall \ell \in A_\ell \quad (8.12b)$$

Na k -ésima iteração, um novo conjunto de multiplicadores de Lagrange é obtido da seguinte forma:

$$(\lambda, \theta)^{k+1} = (\lambda, \theta)^k + t^k \bar{d}_1^k(\bar{w}_\ell^t, \bar{f}_\ell^{pt}) \quad (8.13)$$

em que t^k é um escalar positivo (tamanho do passo) e $\bar{d}_1^k(\bar{w}_\ell^t, \bar{f}_\ell^{pt})$ é uma direção de subida gerada a partir de ξ_1^{TGRP} .

8.2.4 Heurística Lagrangeana para TGRP

Novamente, de forma semelhante ao realizado para o TGP, uma heurística lagrangeana para o TGRP foi implementada e testada. Conforme descrito na seção 5.5, essa heurística representa uma restrição da implementação do algoritmo de *branch-and-bound*, em que apenas o nó inicial (equivalente ao problema original) é explorado.

Para tanto, o TGRP foi relaxado (via relaxação lagrangeana) para o cálculo de limites inferiores conforme já descrito na seção 8.2.1. Implementou-se, então, um método de resolução do problema lagrangeano dual (método de subgradiente) para atualização desses limites como descrito na seção 8.2.3. Além disso, para obtenção de soluções viáveis a heurística descrita na seção 8.2.2 também foi implementada.

8.2.5 Experimentos Computacionais

A heurística lagrangeana descrita na seção 8.2.4 foi implementada em C e testada em uma estação Sun-Blade 100 com processador RISC de 500 MHz e 1 GB RAM. De modo a se resolver o problema de fluxo de custo mínimo dinâmico (ou multiperíodo), optou-se pela utilização do pacote de otimização *CPLEX* (versão 7.0.0) [76]. Para o cálculo de caminhos mínimos, implementou-se o mesmo algoritmo usado anteriormente para codificação da heurística lagrangeana para o TGP – algoritmo de Dijkstra utilizando *D-Heaps* [146].

A heurística lagrangeana para o TGRP foi testada em 60 instâncias agrupadas em dois conjuntos. O conjunto I possui 30 instâncias do TGRP para redes em anel com 5, 10 e 16 nós; enquanto que o conjunto II possui 30 instâncias para redes irregulares com 6, 10 e 15 nós. As topologias das redes irregulares são as mesmas utilizadas nos testes realizados com a heurística lagrangeana para o TGP (ver Figura 5.3). Para cada uma das topologias, padrões de tráfego incrementais e não uniformes – conforme seção 3.4 – para 2 e 3 períodos foram gerados aleatoriamente (5 instâncias para cada combinação de topologia e número de períodos). O primeiro período possuía 40 produtos, enquanto que no segundo período foram adicionados mais 20 produtos (totalizando 60 produtos). No terceiro período (caso existisse), o número total de produtos aumentou para 80, acrescentando-se outros

Tabela 8.1: Resultados do **TGRP** para Redes em Anel (valores médios)

V	N	A	T	P	CPLEX		Heurística Lagr.		
					Iter	Tempo (seg)	Iter	Tempo (seg)	Gap (%)
5	25	40	2	60	3077	15,13	2	0,00	0,00
			3	80	4889	30,70	2	0,00	0,00
10	50	80	2	60	9264	96,36	51	26,65	0,08
			3	80	20093	191,72	23	23,19	0,11
16	80	128	2	60	176543	1847,38	70	73,81	0,22
			3	80	80463	1452,33	89	164,00	0,39

20 produtos aos 60 já existentes no período anterior. Todos os produtos possuíam um valor de demanda unitário (isto é, apenas um canal) e o número máximo de produtos transportados em um *container virtual* em um dado período foi fixado em 63. O número máximo de iterações do método de subgradiente (*Max_Iter*) foi fixado em 200, exceto para as redes irregulares com 15 nós em que se permitiu realizar até 300 iterações do método de subgradiente. Os demais parâmetros foram fixados como na seção 5.6. Além disso, para avaliação dos custos dos períodos futuros adotou-se a mesma função objetivo – equação (3.7) – juntamente com uma taxa de desconto de 10%.

A heurística lagrangeana para o **TGRP** foi capaz de encontrar uma solução ótima ou “quase-ótima” para 58 instâncias (53 soluções ótimas e 5 soluções “quase-ótimas”). Os resultados da heurística foram comparados com aqueles obtidos pelo *CPLEX* - versão 7.0.0 (na mesma estação de trabalho) e são exibidos nas Tabelas 8.1 e 8.2. As colunas dessas tabelas são as mesmas das tabelas apresentadas na seção 5.6 com os resultados da aplicação da heurística lagrangeana para o **TGP**, exceto pela quarta coluna da Tabela 8.1 e da quinta coluna da Tabela 8.2 que apresentam o número de períodos de cada subconjunto de problemas.

Apesar das instâncias utilizadas nos testes possuírem um grande número de variáveis e restrições¹, a heurística lagrangeana se comportou de forma bastante estável e o tempo gasto não foi significativo (exceto para 2 instâncias da rede irregular com 15 nós). Esta situação parece estar associada ao conjunto de inequações válidas (cortes) utilizadas na melhoria do valor do limite inferior – (8.5a) e (8.5b), pois nos dois casos citados, os cortes utilizados não parecem suficientes para promover um aumento significativo no valor do limite inferior (essas instâncias apresentaram *gap de dualidade* $\approx 16\%$).

8.3 Abordagem Lagrangeana do **STGP**

Em vista do que foi apresentado nas subseções 3.2.2, poucos ajustes nas implementações realizadas para resolver o **TGP** são necessários para se resolver a versão limitada desse problema **STGP**, em que a capacidade de realização de *grooming*

¹A formulação para 2 períodos da rede em anel com 16 nós possui 15616 variáveis e 10112 restrições; enquanto que a da rede irregular com 15 nós, 28060 variáveis e 12738 restrições. Já a formulação da rede irregular com 15 nós e 3 períodos possui 55890 variáveis e 24932 restrições.

Tabela 8.2: Resultados do **TGRP** para Redes Irregulares (valores médios)

V	E	N	A	T	P	CPLEX		Heurística Lagr.		
						Iter	Tempo (seg)	Iter	Tempo (seg)	Gap (%)
6	8	38	76	2	60	12833	122,18	39	19,40	0,12
				3	80	35360	362,93	38	35,23	0,13
10	24	58	108	2	60	32557	344,54	79	60,55	0,13
				3	80	54301	635,95	64	88,98	0,14
15	21	99	230	2	60	868717	250980,42	300	6017,50	4,17
				3	80	5605345	168199,46	300	996,10	4,84

está limitada a apenas alguns dos elementos da rede representados pelo subconjunto N^g (ver seção 3.2).

De fato, a principal modificação está relacionada à estrutura da representação estendida e, conseqüentemente, da representação em camadas. Sendo assim, ela não afeta diretamente a formulação condensada de tal problema restrito, isto é, $STGP_C$, nem implica em profundas alterações nos limites inferiores obtidos através da relaxação lagrangeana.

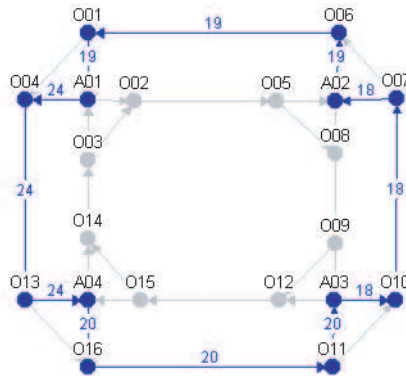
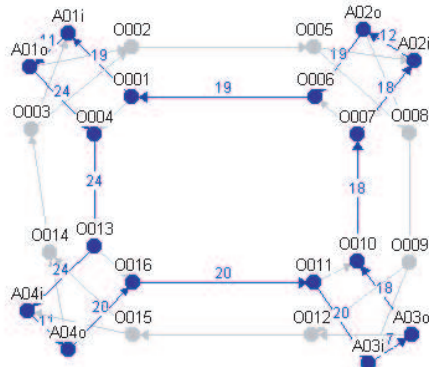
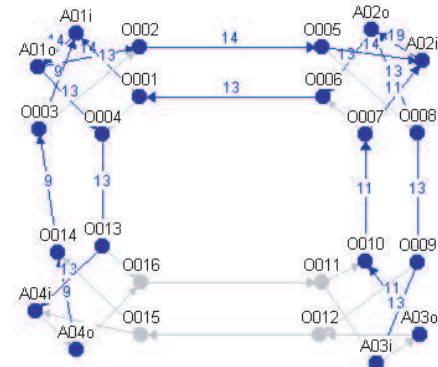
Contudo, deve-se atentar para necessidade da extensão do conjunto de arcos de ambas as representações, de modo a incluir os arcos conectando os nós de *drop* aos respectivos nós de *add* (ambos correspondentes aos elementos da rede pertencentes a N^g). Maiores detalhes são descritos na subseção 3.2.2. Além disso, deve-se estender o conjunto de variáveis de fluxo de modo a se representar a realização de *grooming* (isto é, envio de produtos de um nó de *drop* para um nó de *add*).

Devido ao seu “bom” desempenho nos testes apresentados anteriormente, tais modificações foram realizadas para a implementação da heurística lagrangeana utilizando a relaxação #2 do **TGP**, produzindo, dessa forma, uma heurística lagrangeana para o **STGP**.

A Figura 8.1 apresenta os resultados obtidos pela aplicação de tal heurística a uma rede em anel com 4 nós. A Figura 8.1(a) exibe o resultado obtido pela heurística lagrangeana para o **TGP** (sem restrições à capacidade de realizar *grooming*). Já a Figura 8.1(b) apresenta o resultado obtido pela heurística lagrangeana para o **STGP** quando todos os elementos da rede pertencem a N^g , isto é, $N^g = N^e$ (conforme discutido na seção 3.2). Como era de se esperar, os resultados exibidos pelas Figuras 8.1(a) e 8.1(b) coincidem. Finalmente, a Figura 8.1(c) exibe a solução obtida pela heurística lagrangeana para o **STGP** quando apenas os dois nós da parte superior da figura possuem a capacidade de realizar *grooming*. Deve-se notar que parte do tráfego foi redirecionada de modo a se utilizar a capacidade de *grooming* de tais nós a um custo mínimo.

8.3.1 Experimentos Computacionais

De modo a se avaliar, os resultados pela heurística lagrangeana para o **STGP**, utilizou-se as mesmas instâncias descritas na seção 5.6 (150 problemas divididos em 3 conjuntos) para a realização de experimentos computacionais. Além disso, selecionou-se aleatoriamente um nó qualquer da rede $i \in N^e$ em que não existiria

(a) Solução Obtida para **TGP**(b) Resultado para $N^g = N^e$ (c) Resultado para $N^g = \{1, 2\}$ Figura 8.1: Teste da Heurística Lagrangeana para **STGP**

a capacidade de realizar *grooming*, enquanto que o restante dos elementos da rede, por sua vez, teriam essa capacidade, isto é, $N^{ad} = \{i\}$ e $N^g = N^e \setminus \{i\}$. Nesses experimentos, fixou-se $Max_Iter = 1000$ e o restante dos parâmetros foi definido da mesma maneira que na seção 5.6.

A Tabela 8.3 apresenta os valores médios dos resultados obtidos pela heurística lagrangeana do **STGP** para as redes em anel. A primeira coluna dessa tabela apresenta a identificação do subconjunto de instâncias. Em seguida, os valores médios para o número de iterações, para o tempo gasto (em segundos) e para o *gap de dualidade* são apresentados na segunda, terceira e quarta colunas, respectivamente. Por fim, as três últimas colunas, apresentam os valores médios para o número de multiplexadores *add-drop*, para o número de *lightpaths* e para o número de *wavelength-links* obtidos pela melhor solução encontrada durante o processo de busca.

Deve-se ressaltar que, diferentemente do ocorrido com o **TGP**, a heurística lagrangeana do **STGP** encontrou maiores dificuldades na resolução do problema. Isso pode ser observado através dos altos valores médios para o *gap de dualidade* ao final da aplicação da heurística. Na verdade, ela obteve um *gap* acima de 5% para 26 das 90 instâncias das redes em anel.

Já a Tabela 8.4 apresenta os valores médios dos resultados obtidos pela heurística lagrangeana do **STGP** para as redes irregulares (inclusive para a rede Pan-

Tabela 8.3: Resultados do **STGP** para Redes em Anel (valores médios)

Instância	Iter	Tempo (seg)	Gap (%)	ADMs	LPs	WLs
C04-40	1000	1,65	30,67	7,0	5,8	11,6
C04-80	77	0,30	0,03	4,0	4,0	8,0
C04-120	653	4,64	11,72	5,0	4,4	9,6
C05-40	7	0,02	0,05	5,0	5,0	10,0
C05-80	16	0,09	0,06	5,0	5,0	10,0
C05-120	470	4,42	6,71	5,8	5,4	10,8
C06-40	7	0,02	0,02	6,0	6,0	12,0
C06-80	14	0,09	0,07	6,0	6,0	12,0
C06-120	653	8,08	8,60	7,2	6,6	13,2
C09-40	7	0,04	0,04	9,0	9,0	18,0
C09-80	14	0,16	0,06	9,0	9,0	18,0
C09-120	1000	21,40	14,72	11,6	10,6	23,2
C10-40	206	1,38	1,85	10,4	10,2	20,4
C10-80	19	0,26	0,07	10,0	10,0	20,0
C10-120	640	15,69	9,80	12,0	11,2	22,4
C16-40	1000	12,93	17,19	20,2	18,4	37,8
C16-80	14	0,40	0,04	16,0	16,0	32,0
C16-120	821	38,12	11,10	19,4	18,0	41,8

Européia). As colunas nessa tabela são as mesmas da tabela anterior.

Observa-se que o desempenho da heurística lagrangeana neste caso piora bastante, apresentando valores elevados de *gap de dualidade* para quase todos os subconjuntos das redes irregulares. A situação, na realidade, parece ter se revertido por completo, pois a heurística lagrangeana conseguiu obter a solução ótima (ou “quase-ótima”) apenas para algumas instâncias da rede irregular com 6 nós e baixo ou médio volume de tráfego (40 ou 80 produtos).

Os resultados obtidos anteriormente não parecem indicar que tal situação possa ser contornada eficientemente pela simples adoção de um procedimento de busca (isto é, algoritmo de *branch-and-bound*). Diferentemente do que ocorreu com o **TGRP**, a extensão para o **STGP** das abordagens desenvolvidas para o **TGP** não apresenta bons resultados, levando-nos a questionar sua “real” adequação.

Em um primeiro momento, pode-se atribuir, pelo menos em parte, o baixo desempenho da heurística lagrangeana para o **STGP** a uma eventual inadequação da heurística para geração de limites superiores (semelhante ao que se observou com as abordagens poliédricas). Todavia, não se conseguiu detectar “possíveis falhas” no funcionamento da heurística.

Apesar da aparente similaridade, os resultados parecem indicar que os dois problemas não devem ser tão semelhantes como se imaginava a princípio. Na verdade, se considerarmos que todos os nós da rede possuem capacidade de realizar *grooming* e atribuímos um custo diferente de zero aos arcos conectando os nós de *drop* aos nós de *add*, o **STGP** se torna um problema de localização de facilidades capacitado [107, 108, 123] para o qual se encontra na literatura vários trabalhos que confirmam a existência de um *gap de dualidade* elevado.

Tabela 8.4: Resultados do **STGP** para Redes Irregulares (valores médios)

Instância	Iter	Tempo (seg)	Gap (%)	ADMs	LPs	WLs
M06-40	604	2,51	7,28	7,0	6,4	13,0
M06-80	444	4,02	4,44	6,6	6,2	12,6
M06-120	1000	16,20	16,48	7,8	7,2	15,2
M10-40	1000	7,42	31,20	16,2	14,6	30,4
M10-80	1000	17,09	19,18	14,2	12,4	25,0
M10-120	1000	27,51	17,72	12,8	12,2	24,4
M15-80	1000	23,65	23,41	21,8	20,0	41,6
M15-120	1000	41,45	27,41	23,4	20,8	44,0
N20-80	1000	48,59	27,59	31,2	27,6	58,0
N20-120	1000	64,07	19,15	26,0	24,8	53,2
Euro-80	1000	108,38	24,73	29,2	26,6	53,6
Euro-120	1000	180,53	28,22	31,2	27,8	57,4

8.4 Considerações Finais

Em virtude do relativo sucesso da abordagem lagrangeana na resolução do **TGP**, neste capítulo, investigou-se a adequação dessa abordagem para a resolução do **TGRP** e do **STGP**.

No tocante ao **TGRP**, os resultados são bastante promissores. Após realizadas as adaptações necessárias aos procedimentos de geração de limites inferiores e superiores, bem como no método de subgradiente, os experimentos computacionais indicam que a abordagem lagrangeana parece ser eficaz na resolução do **TGRP**. Deve-se destacar, contudo, que o limite inferior desenvolvido só é válido para as situações em que o tráfego é incremental. Para os demais tipos de tráfego, novos limites inferiores necessitam ser formalizados, implementados e testados.

Já para o **STGP**, apesar da heurística lagrangeana ser capaz de encontrar soluções viáveis para o problema, é necessário uma investigação mais profunda de modo a se determinar de forma mais precisa a real adequação (ou não) da abordagem lagrangeana proposta para o **TGP**. Nesse sentido, os resultados obtidos não permitem, pelo menos até o presente momento, que se afirme ou conteste sua adequação.

Outra dificuldade a ser enfrentada é a obtenção e/ou implementação de soluções alternativas (mesmo que heurísticas) para esses problemas, de modo a se possibilitar uma avaliação comparativa. Infelizmente, a literatura sobre tais problema é ainda escassa.

Por fim, vale mencionar que, em vista das dificuldades já enfrentadas por essas abordagens na tentativa de solucionar o **TGP**, nenhum esforço foi realizado no sentido de se produzir algoritmos para resolução do **TGRP** ou do **STGP** utilizando-se de abordagens poliédricas.

Capítulo 9

Conclusões e Trabalhos Futuros

9.1 Conclusões

Neste trabalho, apresentou-se um estudo profundo sobre o *Traffic Grooming* em redes óticas WDM independentemente da topologia da rede física subjacente. Além disso, realizou-se uma investigação preliminar sobre a adequação dos métodos desenvolvidos na resolução de uma versão do problema em que se considere a reconfiguração da rede ao longo de um horizonte de tempo limitado e de outra, em que apenas alguns dos elementos da rede são capazes de realizar *grooming*.

Em relação ao objetivo geral deste trabalho, deve-se destacar como principal contribuição teórica a proposição e avaliação de uma “nova” formulação natural para o **TGP**, a partir de uma representação estendida para a topologia de rede. Essa formulação se mostrou capaz de capturar todos os detalhes tecnológicos envolvidos no problema evitando, contudo, redundâncias desnecessárias e viabilizando, assim, a resolução do problema. Ainda no âmbito das contribuições teóricas, deve-se mencionar a obtenção de um conjunto de inequações válidas para a formulação proposta, obtidas a partir de um estudo da estrutura facial do poliedro associado ao conjunto de soluções do **TGP**. Tal estudo, apesar de não ter sido capaz de demonstrar que tais inequações válidas representam facetas do mencionado poliedro, conseguiu produzir *planos-de-corte* de “boa” qualidade.

Já em relação as contribuições práticas, merece destaque a proposição e utilização de uma representação em camadas para a topologia de rede cuja adoção permitiu uma simplificação da formulação apresentada para o **TGP**, possibilitando, assim, relacioná-lo a outros problemas encontrados na literatura. Essa formulação “simplificada” também foi utilizada como base para todos os métodos de resolução desenvolvidos e testados neste trabalho.

No tocante às abordagens utilizadas para o desenvolvimento de métodos de resolução do **TGP**, os experimentos computacionais apontam para superioridade da abordagem lagrangeana e, em particular, da heurística lagrangeana, em relação às demais abordagens desenvolvidas e testadas neste trabalho. Mais ainda, constatou-se empiricamente que o procedimento de geração de limites inferiores foi mais eficaz quando se utilizou a relaxação lagrangeana em que manteve a maior quantidade de informações sobre a estrutura do problema (isto é, a relaxação #2). Verificou-se,

também, que “boas” soluções poderiam ser obtidas a partir da restrição do processo de busca apenas à exploração do problema original (o que equivale, à aplicação apenas da heurística lagrangeana) ou, ainda, da aplicação de uma forma limitada do mesmo. Na realidade, a comparação entre os resultados obtidos pela heurística lagrangeana e os resultados dos algoritmos de *branch-and-bound* demonstrou a eficácia da heurística lagrangeana na obtenção de “boas” soluções para o problema, utilizando, contudo, apenas uma pequena fração do tempo necessário aos algoritmos de *branch-and-bound*.

Já os resultados obtidos pela abordagens poliédricas apontam, pelo menos inicialmente, para uma superioridade da abordagem lagrangeana, tendo em vista que ambas as abordagens – poliédrica e lagrangeana – obtêm soluções de qualidade equivalente, entretanto as implementações dos métodos baseados na abordagem lagrangeana apresentam um desempenho muito melhor (isto é, consumo de tempo menor). Deve-se, ainda, mencionar o fato das abordagens poliédricas não conseguirem lidar adequadamente com as grandes instâncias do problema, em virtude do elevado número de variáveis e restrições presentes quer na formulação original quer nas reformulações geradas a partir da adição dos cortes, além das dificuldades encontradas na geração de limites superiores (soluções viáveis) a partir das soluções obtidas para a relaxação linear da formulação do problema.

Em virtude do relativo sucesso da abordagem lagrangeana na resolução do **TGP**, também se investigou a adequação dessa abordagem para a resolução do **TGRP** e do **STGP**. No tocante ao **TGRP**, os resultados são bastante promissores e os experimentos computacionais indicam que a abordagem lagrangeana parece ser eficaz na resolução do **TGRP**. Entretanto, deve-se destacar que o procedimento para geração de limites inferiores desenvolvido só é válido para as situações em que o tráfego é incremental. Já em relação ao **STGP**, apesar da heurística lagrangeana ser capaz de encontrar soluções viáveis para o problema, é necessário uma investigação mais profunda de modo a se determinar de forma mais precisa a real adequação (ou não) da abordagem lagrangeana proposta para o **TGP**. Nesse sentido, os resultados obtidos não permitem, pelo menos até o presente momento, que se afirme ou conteste sua adequação. Além disso, outra dificuldade a ser enfrentada é a obtenção e/ou implementação de soluções alternativas (mesmo que heurísticas) para esses problemas, de modo a se possibilitar uma avaliação comparativa. Infelizmente, a literatura sobre tais problema é ainda escassa.

9.2 Trabalhos Futuros

Em face do escopo abrangente do presente trabalho, diversas linhas distintas de investigação poderiam ser exploradas no futuro. Contudo, optou-se por listar a seguir aquelas que aparentemente são mais promissoras ou que se destacam por sua relevância.

Em primeiro lugar, poder-se-ia investigar mecanismos que permitissem melhorar o desempenho das abordagens poliédricas desenvolvidas para o **TGP**. Isto, em verdade, corresponde a duas linhas distintas de atuação. A primeira delas seria voltada para a melhoria do procedimento de geração de limites superiores a partir das soluções obtidas para a relaxação linear; enquanto que uma segunda linha de atuação

estaria relacionada ao aprofundamento do estudo da estrutura facial do poliedro associado ao conjunto de soluções do **TGP**, de modo a se obter outras famílias de inequações válidas (cortes) que se mostrem, eventualmente, mais eficazes em lidar com as grandes instâncias do problema. Em verdade, apesar da primeira delas parecer mais promissora, uma investigação profunda da estrutura facial do poliedro mencionada poderia vir a produzir resultados de maior relevância.

Já em relação às versões modificadas do problema, **TGRP** e **STGP**, as possíveis linhas de atuação no futuro são inúmeras. No tocante ao **TGRP**, deve-se investigar a implementação de novos limites inferiores para padrões de tráfego diferentes daquele focado neste trabalho (apenas tráfego incremental). A adaptação dos métodos de resolução, especialmente, da heurística lagrangeana, para situações mais dinâmicas e, até mesmo, “on-line” constitui outra linha de investigação interessante. Já em relação ao **STGP**, deve-se aprofundar as análises realizadas de modo a se determinar a real eficácia da extensão das abordagens propostas neste trabalho para sua resolução. Além disso, seu relacionamento com o problema de localização de facilidades capacitado também necessita ser melhor investigado.

Por fim, uma outra linha de investigação poderia ser dedicada à avaliação da extensão das abordagens propostas neste trabalho para o **TGP** em redes óticas *multicast* (isto é, aquelas em que o tráfego possui uma origem e vários destinos). Esse problema representa, em verdade, uma outra extensão natural do problema em que o roteamento de tráfego ponto-a-ponto é substituído pelo roteamento *multicast*. Além disso, acredita-se que sua resolução se tornará cada vez mais necessária no futuro.

Referências Bibliográficas

- [1] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin. Network flows. In G.L. Nemhauser, A.H.G. Rinnooy Kan, and M.J. Todd, editors, *Handbooks in Operations Research and Management Science*, volume 1, pages 211–369. Elsevier Science Publishers B.V., 1989.
- [2] A. Atamtürk. On capacitated network design cut-set polyhedra. *Mathematical Programming*, 92(3):425–437, 2002.
- [3] A. Atamtürk and D. Rajan. On splittable and unsplittable flow capacitated network design arc-set polyhedra. *Mathematical Programming*, 92(2):315–333, 2002.
- [4] D. Avis. On extreme rays of the metric cone. *Can. J. Math.*, 32:126–144, 1980.
- [5] E. Balas, S. Ceria, and G. Cornuéjols. A lift-and-project cutting plane algorithm for mixed 0-1 programs. *Mathematical Programming*, 58:295–324, 1993.
- [6] D. Banerjee and B. Mukherjee. A practical approach for routing and wavelength assignment in large wavelength-routed optical networks. *IEEE JSAC*, 14(5):903–908, 1996.
- [7] D. Banerjee and B. Mukherjee. Wavelength-routed optical networks: linear formulation, resource budgeting tradeoffs, and a reconfiguration study. In *Proc. IEEE INFOCOM*, pages 269–276, 1997.
- [8] S. Banerjee. *Optimally-structured high-speed metropolitan area networks with distributed control*. PhD thesis, Dep. Computer Science, University of California, Davis, CA, 1992.
- [9] S. Banerjee, V. Jain, and S. Shah. Regular multihop logical topologies for lightwave networks. *IEEE Communications Surveys*, pages 2–18, 1999.
- [10] S. Banerjee, J. Yoo, and C. Chen. Design of wavelength-routed optical networks for packet switched traffic. *Journal of Lightwave Technology*, 15(9):1636–1646, 1997.
- [11] F. Barahona. Network design using cut inequalities. *SIAM J. Optimization*, 6(3):823–837, 1996.
- [12] F. Barahona and A.R. Mahjoub. On the cut polytope. *Mathematical Programming*, 36:157–173, 1986.

- [13] C. Barnhart. Dual-ascent methods for large-scale multicommodity flow problems. *Naval Research Logistics*, 40:305–324, 1993.
- [14] R.S. Barr and R. A. Patterson. Grooming telecommunication networks. *Optical Networks Magazine*, 2(3):20–23, May/June 2001.
- [15] M.S. Bazaraa and H.D. Sherali. On the choice of step size in subgradient optimization. *European Journal of Operational Research*, 7:380–388, 1981.
- [16] J.E. Beasley. An algorithm for solving large capacitated warehouse location problems. *Journal of the Operational Research*, 33:314–325, 1988.
- [17] A. Bellato, M. Fontana, G. Gasparini, G. Grammel, J. Jones, and D. Papadimitriou. Traffic engineering routing extensions to OSPF and ISIS for GMPLS control of G.709 optical transport networks. <http://search.ietf.org/internet-rafts/draft-gaspariniccamp-gmpls-g709-ospf-isis-00.txt>, November 2001.
- [18] R. Berry and E. Modiano. Reducing electronic multiplexing costs in SONET/WDM rings with dynamically changing traffic. *IEEE Journal on Selected Areas in Communications*, 18(10):1961–1971, October 2000.
- [19] D.P. Bertsekas and P. Tseng. RELAX-IV: a faster version of the RELAX code for solving minimum cost flow problems. Technical report, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1994.
- [20] D. Bienstock, S. Chopra, O. Günluk, and C.-Y. Tsai. Minimum cost capacity installation for multicommodity networks. *Mathematical Programming*, 81:177–199, 1998.
- [21] D. Bienstock and O. Günluk. Capacitated network design - polyhedral structure and computation. *INFORMS Journal of Computing*, 8:243–259, 1996.
- [22] N. Bitar, R. Rao, and K. Muthukrishnan. Traffic engineering extensions to OSPF. <http://search.ietf.org/internet-drafts/draft-bitar-rao-ospf-diffserv-mpls-01.txt>, July 2001.
- [23] M.S. Borella, J.P. Jue, D. Banerjee, B. Ramamurthy, and B. Mukherjee. Optical components for WDM lightwave networks. *Proceedings of the IEEE*, 85(8):1274–1307, 1997.
- [24] E.A. Boyd. Fenchel cutting planes for integer programs. *Operations Research*, 42:53–64, 1994.
- [25] R. Busacker and P. Gowen. A procedure for determining a family of minimum-cost network flow patterns. ORO Technical Report 15, John Hopkins University, 1961.
- [26] P.M. Camerini, L. Fratta, and F. Maffioli. On improving relaxation methods by modified gradient techniques. *Math. Programming Study*, 3:26–34, 1975.

- [27] J. Castro and N. Nabona. An implementation of linear and nonlinear multi-commodity network flows. *European Journal of Operational Research*, 92:37–53, 1996.
- [28] C. Chen and S. Banerjee. Optical switch configuration and lightpath assignment in wavelength routing multihop lightwave networks. In *Proc. IEEE INFOCOM*, pages 1300–1307, 1995.
- [29] E.W. Cheney and A.A. Goldstein. Newton’s method for convex programming and Tchebycheff approximation. *Numerische Mathematik*, 1:253–268, 1959.
- [30] A.L. Chiu and E.H. Modiano. Traffic grooming in algorithms for reducing electronic multiplexing costs in WDM ring networks. *IEEE/OSA Journal of Lightwave Technology*, 18(1):2–12, January 2000.
- [31] I. Chlamtac, A. Ganz, and G. Karmi. Lightpath communications: an approach to high bandwidth optical WANs. *IEEE Trans. Communications*, 40(7):1171–1182, 1992.
- [32] I. Chlamtac, A. Ganz, and G. Karmi. Lightnets: topologies for high-speed optical networks. *Journal of Lightwave Technology*, 11(5/6):951–961, 1993.
- [33] W. Cho, J. Wang, and B. Mukherjee. Improved approaches for cost-effective traffic grooming in WDM ring networks: Uniform-traffic case. *Photon. Network Commun.*, 3(2):245–254, Apr. 2001.
- [34] T. Cinkler, D. Marx, C.P. Larsen, and D. Fogaras. Heuristic algorithms for joint configuration of optical and electrical layer in multi-hop wavelength routing networks. In *Proc. IEEE INFOCOM*, 2000.
- [35] S. Corteel, D. Gardy, M. Valencia-Pabon, D. Barth, and A. Denise. The permutation-path coloring problem on trees. *Theoretical Computer Science*, 297:119–143, 2003.
- [36] L.A. Cox Jr. and J.R. Sanchez. Cost savings from optimized packing and grooming of optical circuits: mesh versus ring comparisons. *Optical Networks Magazine*, 2(3):72–90, May/June 2001.
- [37] H. Crowder. Computational improvements for subgradient optimization. In *Symposia Mathematica*, volume XIX, Academic Press, London, 1976.
- [38] H. Crowder, E.L. Johnson, and M.W. Padberg. Solving large-scale zero-one linear programming problems. *Operations Research*, 31:803–834, 1983.
- [39] R.J. Dakin. A tree search algorithm for mixed integer programming problems. *Computer Journal*, 8:250–255, 1965.
- [40] G.B. Dantzig, R. Fulkerson, and S.M. Johnson. Solution of a large scale traveling salesman problem. *Operations Research*, 2:393–410, 1954.

- [41] S.S. Dixit and Antti Ylä-Jääski. WDM optical networks: a reality check. *IEEE Communications Magazine*, 38(3):58–60, 2000.
- [42] R. Dutta. *Virtual Topology Design for Traffic Grooming in WDM Networks*. PhD thesis, North Carolina State University, Raleigh, NC, Aug. 2001.
- [43] R. Dutta and G.N. Rouskas. A survey of virtual topology design algorithms for wavelength routed optical networks. *Optical Networks Magazine*, 1(1):73–89, 2000.
- [44] R. Dutta and G.N. Rouskas. On optimal traffic grooming in WDM rings. *IEEE Journal on Selected Areas in Communications*, 20(1):110–121, Jan. 2002.
- [45] R. Dutta and G.N. Rouskas. Traffic grooming in WDM networks: Past and future. *IEEE Network*, pages 46–56, Nov./Dec. 2002.
- [46] J. Edmonds. Maximum matching and a polyhedron with 0,1-vertices. *Journal of Research of the National Bureau of Standards (B)*, 69:125–130, 1965.
- [47] C.E. Ferreira. *On Combinatorial Optimization Problems Arising in Computer System Design*. PhD thesis, Technische Universität Berlin, Berlin, Germany, 1994.
- [48] C.E. Ferreira and Y. Wakabayashi. *Combinatória Poliédrica e Planos-de-Corte Faciais*. Instituto de Computação, UNICAMP, Campinas, 1996.
- [49] M.L. Fisher. The Lagrangean relaxation method for solving integer programming problems. *Management Science*, 27:1–18, 1981.
- [50] M.L. Fisher. An application oriented guide to Lagrangean relaxation. *Interfaces*, 15:10–21, 1985.
- [51] L. Ford and D. Fulkerson. *Flows in Networks*. Princeton University Press, Princeton, NJ, 1962.
- [52] J.-M. Gauthier and G. Ribière. Experiments in mixed-integer linear programming using pseudo costs. *Mathematical Programming*, 12:26–47, 1977.
- [53] A.M. Geoffrion. Lagrangean relaxation for integer programming. *Mathematical Programming Study*, 2:82–114, 1974.
- [54] O. Gerstel, P. Lin, and G. Sasaki. Wavelength assignment in WDM rings to minimize cost of embedded SONET rings. In *Proc. IEEE INFOCOM*, volume 1, pages 94–101, San Francisco, CA, March 1998.
- [55] O. Gerstel, R. Ramaswami, and G.H. Sasaki. Cost-effective traffic grooming in WDM rings. *IEEE/ACM Trans. Networking*, 8(5):618–630, October 2000.
- [56] J.L. Goffin. On convergence rates of subgradient optimization methods. *Mathematical Programming*, 13:329–347, 1977.

- [57] J.L. Goffin and J.P. Vial. Convex nondifferentiable optimization: an interior point perspective. In *2nd School on Optimization*, Universidad de Chile, Santiago, Chile, 1998.
- [58] R.E. Gomory. Outline on an algorithm for integer solutions to linear programs. *Bulletin of American Mathematical Society*, 64:275–278, 1958.
- [59] R.E. Gomory. Solving linear programming problems in integers. In *Proceedings of the Symposium of Applied Mathematics 10*, pages 211–215, 1960.
- [60] R.E. Gomory. An algorithm for integer solutions to linear programs. In R.L. Graves and P. Wolfe, editors, *Recent Advances in Mathematical Programming*, pages 269–302. McGraw Hill, New York, 1963.
- [61] M. Grötschel. *Polyedrische Charakterisierungen Kombinatorischer Optimierungsprobleme*. PhD thesis, Verlag A. Hain, Meisenheim, 1977.
- [62] M. Grötschel, M. Jünger, and G. Reinelt. A cutting plane algorithm for the linear ordering problem. *Operations Research*, 32:1195–1220, 1985.
- [63] M. Grötschel, M. Jünger, and G. Reinelt. Facets of the linear ordering polytope. *Mathematical Programming*, 33:43–60, 1985.
- [64] M. Grötschel and M.W. Padberg. Polyhedral aspects of the traveling salesman problem I : Theory. In E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, and D.B. Shmoys, editors, *The Traveling Salesman Problem*. Wiley, New York, 1985.
- [65] M. Grötschel and M.W. Padberg. Polyhedral theory. In E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, and D.B. Shmoys, editors, *The Traveling Salesman Problem*. Wiley, New York, 1985.
- [66] M. Grötschel and Y. Wakabayashi. A cutting plane algorithm for a clustering problem. *Mathematical Programming*, 45:59–96, 1989.
- [67] M. Grötschel and Y. Wakabayashi. Facets for the clique partitioning polytope. *Mathematical Programming*, 47:367–388, 1990.
- [68] U. Gupta, D. Lee, and J. Leung. Efficient algorithms for interval graphs and circular-arc graphs. *Networks*, 12:459–467, 1982.
- [69] E. Harder, S.-K. Lee, and H.-A. Choi. On wavelength assignment in WDM optical networks. In *Proc. on the 4th International Conference on Massively Parallel Processing Using Optical Interconnections*, Geneve, Switzerland, April 1997.
- [70] M. Held, P. Wolfe, and H.D. Crowder. Validation of subgradient optimization. *Mathematical Programming*, 6:62–88, 1974.
- [71] J.B. Hiriart and C. Lemaréchal. *Convex Analysis and Minimization Algorithms*. Springer-Verlag, 1991.

- [72] K. Hoffman and M.W. Padberg. Improving LP-representations of zero-one linear programs for branch and cut. *ORSA Journal on Computing*, 3:121–134, 1991.
- [73] K. Holmberg. Lagrangean heuristics for linear cost multicommodity network flow problems. Working paper LiTH-MAT/OPT-WP-1995-01, Division of Optimization, Department of Mathematics, Linköping Institute of Technology, Sweden, 1995.
- [74] K. Holmberg and D. Yuan. A Lagrangean approach to network design problems. *Int. Trans. Opl. Res.*, 5(6):529–539, 1998.
- [75] T. Hu. Multi-commodity network flows. *Operations Research*, 11:344–360, 1963.
- [76] ILOG Inc. CPLEX. <http://www.ilog.com>. Mountain View, CA.
- [77] J. Iness. *Efficient use of optical components in WDM-based optical networks*. PhD thesis, Dep. Computer Science, University of California, Davis, CA, 1997.
- [78] INFORMS. COIN-OR: Common Optimization INterface for Operations Research. <http://www.coin-or.org>.
- [79] M. Iri. A new method for solving transportation-network problems. *Journal of the Oper. Res. Society of Japan*, 3:27–87, 1960.
- [80] M. Iri. On an extension of the maximum-flow minimum-cut theorem to multicommodity flows. *Journal of the Oper. Res. Society of Japan*, 13:129–135, 1971.
- [81] W. Jewell. Optimal flow through network. Technical Report 8, OR Center, MIT, Cambridge, MA, 1958.
- [82] M. Jünger. Polyhedral combinatorics and the acyclic subdigraph problem. Research and Exposition in Mathematics 7, Heldermann-Verlag, Berlin, 1985.
- [83] M. Jünger, G. Reinelt, and S. Thienel. Practical problem solving with cutting plane algorithms in combinatorial optimization. Preprint 94-24, Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Universität Heidelberg, März 1994.
- [84] K. Kaklamanis, G. Persiano, T. Erlebach, and K. Jansen. Constrained bipartite edge coloring with applications to wavelength routing. In *Proc. on the 24th International Colloquium on Automata, Languages, and Programming (ICALP)*, Bologna, Italy, July 1997.
- [85] O. Kakusho and K. Onaga. On feasibility conditions of multicommodity flows in networks. *Transactions on Circuit Theory*, 18:425–429, 1971.
- [86] N. Karmakar. A new polynomial-time algorithm for linear programming. *Combinatorica*, pages 373–395, 1984.

- [87] R.M. Karp and C.H. Papadimitriou. On linear characterizations on combinatorial optimization problems. *SIAM Journal on Computing*, 11:620–632, 1982.
- [88] D. Katz, D. Yueng, and K. Kompella. Traffic engineering extensions to OSPF. <http://search.ietf.org/internet-drafts/draft-katz-yeung-ospf-traffic-06.txt>, October 2001.
- [89] J.E. Kelley. The cutting plane method for solving convex programs. *Journal of the SIAM*, 8:703–712, 1960.
- [90] L.G. Khachiyan. A polynomial algorithm in linear programming. *Doklady Akad. Nauk SSSR*, 244:1093–1096, 1979. (Em Russo). (Tradução para Inglês: *Soviet Mathematics Doklady*, 20: 191–194, 1979).
- [91] R. Krishnaswamy and K. Sivaraman. Of logical topologies: a linear formulation for wavelength routed optical networks with no wavelength changers. In *Proc. IEEE INFOCOM*, pages 919–927, 1998.
- [92] J.-F. P. Labourdette. Traffic optimization and reconfiguration management of multiwavelength multihop broadcast lightwave networks. *Computer Networks and ISDN Systems*, 30:981–998, 1998.
- [93] A.H. Land and A.G. Doig. An automatic method for solving discrete programming problems. *Econometrica*, 28:493–520, 1960.
- [94] J.P. Lang et al. Link management protocol (LMP). <http://www.ietf.org/internet-drafts/draft-ietf-ccamp-lmp-09.txt>, June 2003.
- [95] A. Lardies, R. Gupta, and R.A. Patterson. Traffic grooming in a multi-layer network. *Optical Networks Magazine*, 2(3):91–99, May/June 2001.
- [96] C. Lemaréchal. Nondifferentiable optimization. In G.L. Nemhauser, A.H.G. Rinnooy Kan, and M.J. Todd, editors, *Handbooks in Operations Research and Management Science*, volume 1, pages 529–572. Elsevier Science Publishers B.V., 1989.
- [97] A.Y. Levin. On an algorithm for minimizing a convex function. *Soviet Mathematics Doklady*, 6(1):286–290, 1965.
- [98] G. Li and R. Simha. On the wavelength assignment problem in multifiber WDM star and ring networks. *IEEE/ACM Trans. Networking*, 9(1):60–68, February 2001.
- [99] M.V. Lomonosov. Combinatorial approaches to multiflow problems. *Discrete Applied Math.*, 11:1–93, 1985.
- [100] T.L. Magnanti and P. Mirchandani. Shortest path, single origin-destination network design and associated polyhedra. *Networks*, 23:103–121, 1993.

- [101] T.L. Magnanti, P. Mirchandani, and R. Vachani. The convex hull of two core capacitated network design problems. *Mathematical Programming*, 60:233–250, 1993.
- [102] T.L. Magnanti, P. Mirchandani, and R. Vachani. Modeling and solving the two-facility capacitated network loading problem. *Operations Research*, 43(1):142–157, 1995.
- [103] H. Marchand and A. Wolsey. Aggregation and mixed integer rounding to solve mips. *Operations Research*, 49(3):363–371, 2001.
- [104] L. Margara and J. Simon. Wavelength assignment problem on all-optical networks with k fibers per link. In *Proc. on the 27th International Colloquium on Automata, Languages, and Programming (ICALP)*, 2000.
- [105] M. Marsan, A. Bianco, E. Leonardi, and F. Neri. Topologies for wavelength-routing all-optical networks. *IEEE/ACM Trans. Networking*, 1(5):534–546, 1993.
- [106] G.R. Mateus and Z.K.G. Patrocínio Jr. Um algoritmo para a problema de localização capacitado baseado no modelo de fluxo de custo mínimo (in portuguese). Technical Report RT003-91, Depart. Computer Science, Federal University of Minas Gerais, Belo Horizonte, Brasil, 1991.
- [107] G.R. Mateus and Z.K.G. Patrocínio Jr. An algorithm for large scale capacitated location problem in networks. In *IFAC/IFORS/IMACS Symposium Large Scale Systems: Theory and Applications*, pages 65–72, Beijing, P.R. China, August 1992.
- [108] G.R. Mateus and Z.K.G. Patrocínio Jr. Algorithms for capacitated network location problems. In *13th Triennial Conference on Operations Research*, volume 35, Lisbon, Portugal, July 1993.
- [109] G.R. Mateus and Z.K.G. Patrocínio Jr. Um algoritmo de separação e avaliação para um problema de planejamento de redes. In *XXI Seminário Integrado de Software e Hardware*, pages 411–425, Caxambu, Brasil, Julho 1994.
- [110] P.P. Mitra and J.B. Stark. Nonlinear limits to the information capacity of optical fibre communications. *Nature*, 411:1027–1030, 2001.
- [111] B. Mukherjee. *Optical Communication Networks*. McGraw-Hill, New York, 1997.
- [112] B. Mukherjee, S. Ramamurthy, D. Banerjee, and A. Mukherjee. Some principles for designing a wide-area optical network. In *Proc. IEEE INFOCOM*, pages 110–119, 1994.
- [113] B. Mukherjee, S. Ramamurthy, D. Banerjee, and A. Mukherjee. Some principles for designing a wide-area optical network. *IEEE/ACM Trans. Networking*, 4(5):684–696, 1996.

- [114] G.L. Nemhauser and L.A. Wolsey. *Integer and Combinatorial Optimization*. John Wiley & Sons, New York, 1988.
- [115] G.L. Nemhauser and L.A. Wolsey. Integer programming. In G.L. Nemhauser, A.H.G. Rinnooy Kan, and M.J. Todd, editors, *Handbooks in Operations Research and Management Science*, volume 1, pages 447–527. Elsevier Science Publishers B.V., 1989.
- [116] M.W. Padberg and M. Grötschel. Polyhedral aspects of the traveling salesman problem II : Computations. In E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, and D.B. Shmoys, editors, *The Traveling Salesman Problem*. Wiley, New York, 1985.
- [117] M.W. Padberg and G. Rinaldi. Optimization of a 532-city symmetric traveling salesman problem by branch and cut. *Operations Research Letters*, 6:1–7, 1987.
- [118] C.H. Papadimitriou and M. Yannakakis. The complexity of facets (and some facets of complexity). *Journal of Computer and System Science*, 28:244–259, 1984.
- [119] Z.K.G. Patrocínio Jr. and G.R. Mateus. An algorithm for traffic grooming in WDM optical networks. In *Proceedings of IV ALIO/EURO Workshop on Applied Combinatorial Optimization*, pages 117–118, Pucón, Chile, 2002.
- [120] Z.K.G. Patrocínio Jr. and G.R. Mateus. A lagrangian-based heuristic for traffic grooming in WDM optical networks. In *Proceedings of IEEE GLOBECOM 2003 - Global Telecommunication Conference*, pages 2767–2771, San Francisco, USA, 2003.
- [121] Z.K.G. Patrocínio Jr. and G.R. Mateus. A polyhedral approach to traffic grooming problem in WDM optical networks. In *Proceedings of Mathematical Programming in Rio*, pages 94–99, Búzios, Brasil, 2003.
- [122] Z.K.G. Patrocínio Jr., P.P.R. Teixeira Filho, and G.R. Mateus. Traffic grooming and reconfiguration for incremental traffic in WDM optical networks. In *Proceedings of INFORMS INOC 2003 - International Network Optimization Conference*, pages 454–459, Evry/Paris, França, 2003.
- [123] Z.K.G. Patrocínio Jr. Um algoritmo exato para o problema de localização capacitado baseado no modelo de fluxo de custo mínimo (in portuguese). Master's thesis, Depart. Computer Science, Federal University of Minas Gerais, Belo Horizonte, Brasil, Março 1993.
- [124] Z.K.G. Patrocínio Jr and G.R. Mateus. An algorithm for the capacitated location problem based on the minimal cost network flow problem. In *TIMS XXX - SOBRAPO XXIII*, Rio de Janeiro, Brazil, July 1991.
- [125] B.T. Polyak. Minimization of nonsmooth functionals. *USSR Comput. Math. and Math. Phys.*, 9:14–29, 1969.

- [126] W.R. Pulleyblank. Polyhedral combinatorics. In G.L. Nemhauser, A.H.G. Rinnooy Kan, and M.J. Todd, editors, *Handbooks in Operations Research and Management Science*, volume 1, pages 371–446. Elsevier Science Publishers B.V., 1989.
- [127] R. Rado. Theorems on linear combinatorial topology and general measure. *Annals of Mathematics*, 44:228–270, 1943.
- [128] L.A. Raggi. *Métodos de Solução para o Problema de Localização Não Capacitado*. PhD thesis, COPPE, Federal University of Rio de Janeiro, Rio de Janeiro, Brasil, 1987.
- [129] T. Ralphs and L. Ladanyi. SYMPHONY: Single- or Multi-Process Optimization over Networks. <http://branchandcut.org/SYMPHONY>.
- [130] B. Ramamurthy. *Efficient design of wavelength division multiplexing based optical networks*. PhD thesis, Dep. Computer Science, University of California, Davis, CA, 1998.
- [131] R. Ramaswami and G.H. Sasaki. Multiwavelength optical networks with limited wavelength conversion. *IEEE/ACM Trans. Networking*, 6(6):744–754, December 1998.
- [132] R. Ramaswami and K. Sivarajan. Design of logical topologies for wavelength-routed optical networks. *IEEE JSAC*, 14(5):840–851, 1996.
- [133] R. Ramaswami and K.N. Sivarajan. *Optical Networks: A Practical Perspective*. Morgan Kaufmann Publisher Inc., San Francisco, 1998.
- [134] RHK Inc. Network Traffic & Revenue Analysis. Technical Report #30001, RHK Inc., South San Francisco, CA, Oct. 2002.
- [135] RHK Inc. Optical Networks: Global. Technical Report #17100, RHK Inc., South San Francisco, CA, Oct. 2002.
- [136] RHK Inc. Optical Components. Technical Report #33010, RHK Inc., South San Francisco, CA, Mar. 2003.
- [137] RHK Inc. Optical Networks: Europe. Technical Report #30310, RHK Inc., South San Francisco, CA, Mar. 2003.
- [138] RHK Inc. Optical Networks: Global. Technical Report #30005, RHK Inc., South San Francisco, CA, May 2003.
- [139] L. Sahasrabuddhe and B. Mukherjee. Light-trees: optical multicasting for improved performance in wavelength-routed networks. *IEEE Communications Magazine*, 37(2):67–73, February 1999.
- [140] C. Shannon. A mathematical theory of communications. *Bell System Tech. Journal*, 27:379–423, 1948.

- [141] N. Shor. *Minimization Methods for Nondifferentiable Functions*. Springer-Verlag, Berlin, 1985.
- [142] J.M. Simmons, E.L. Goldstein, and A.A.M. Saleh. Quantifying the benefit of wavelength add-drop in WDM rings with distance-independent and dependent traffic. *IEEE/OSA Journal of Lightwave Technology*, 17(1):48–57, January 1999.
- [143] N. Singhal and B. Mukherjee. Architectures and algorithms for multicasting in optical WDM mesh networks using opaque and transparent optical cross-connects. In *Proc. IEEE OFC*, volume 2, Anaheim, CA, March 2001.
- [144] M. Stoer and G. Dahl. A polyhedral approach to multicommodity survivable network design. *Numerische Mathematik*, 68:149–167, 1994.
- [145] A.S. Tanenbaum. *Redes de Computadores*. Campus, Rio de Janeiro, 3 edition, 1997.
- [146] R.E. Tarjan. *Data Structures and Network Algorithms*, volume 44 of *Regional Conference Series in Applied Mathematics*. SIAM, Philadelphia, Pennsylvania, 1983.
- [147] P.P.R. Teixeira Filho, Z.K.G. Patrocínio Jr., and G.R. Mateus. Um algoritmo para reconfiguração e agrupamento de tráfego em redes Óticas WDM. In *Anais do XXXV SBPO - Simpósio Brasileiro de Pesquisa Operacional*, Natal, Brasil, 2003.
- [148] S. Thiagarajan and A. K. Somani. A capacity correlation model for WDM networks with constrained grooming capabilities. In *Proc. IEEE ICC*, Helsinki, Finland, June 2001.
- [149] S. Thiagarajan and A.K. Somani. Capacity fairness of WDM networks with grooming capabilities. *Optical Networks Magazine*, 2(3):24–31, May/June 2001.
- [150] S. Thienel. ABACUS: A Branch And CUt System. http://www.informatik.uni-koeln.de/lj_juenger/projects/abacus.html.
- [151] S. Téral. Optical networking: After the crash. <http://www.rhk.com>, 2002.
- [152] S.P.M. van Hoesel, A.M.C.A. Koster, R.L.M.J. van de Leensel, and M.W.P. Savelsberg. Polyhedral results for the edge capacity polytope. *Mathematical Programming*, 92(2):335–358, 2002.
- [153] T.J. Van Roy and L.A. Wolsey. Solving mixed integer programming problems using automatic reformulation. *Operations Research*, 35:45–57, 1987.
- [154] P.J. Wan, G. Calinescu, L. Liu, and O. Frieder. Grooming of arbitrary traffic in SONET/WDM BLSRs. *IEEE Journal on Selected Areas in Communications*, 18(10):1995–2003, October 2000.

- [155] J. Wang and B. Mukherjee. Interconnected WDM ring networks: Strategies for interconnection and traffic grooming. *Optical Networks Magazine*, 3(5):10–20, Sept./Oct. 2002.
- [156] J. Wang, V.R. Vemuri, W. Cho, and B. Mukherjee. Improved approaches for cost-effective traffic grooming in WDM ring networks: ILP formulations and single-hop and multihop connections. *IEEE/OSA Journal of Lightwave Technology*, 19(11):1645–1653, November 2001.
- [157] Roland Wessály. *Dimensioning Survivable Capacitated NETWORKS*. PhD thesis, Fakultät für Mathematik, Technische Universität Berlin, Berlin, Germany, 2000.
- [158] D. West. *Introduction to Graph Theory*. Prentice Hall, 1996.
- [159] G. Wilfong and P. Winkler. Ring routing and wavelength translation. In *Proc. on the 9th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 333–341, 1998.
- [160] D.G. Wilson and B.D. Rudin. Introduction to the IBM optimization subroutine library. *IBM Systems Journal*, 31(1):4–10, 1992.
- [161] Y. Xu et al. A framework for generalized multi-protocol label switching (GMPLS). <http://search.ietf.org/internet-drafts/draft-many-ccamp-gmpls-framework-00.txt>, September 2001.
- [162] Y. Xu, S.-C. Xu, and B.-X. Wu. Traffic grooming in unidirectional WDM ring networks using genetic algorithms. *Computer Communications*, 25:1185–1194, 2002.
- [163] Y. Xu, S.-C. Xu, and B.-X. Wu. Strictly nonblocking grooming of dynamic traffic in unidirectional SONET/WDM rings using genetic algorithms. *Computer Networks*, 41:227–245, 2003.
- [164] X. Zhang and C. Qiao. An effective and comprehensive approach for traffic grooming and wavelength assignment in SONET/WDM rings. *IEEE/ACM Trans. Networking*, 8(5):608–617, October 2000.
- [165] Z. Zhang and A. Acampora. An heuristic wavelength assignment algorithm for multihop WDM networks with wavelength routing and wavelength re-use. *IEEE/ACM Trans. Networking*, 3(3):281–288, 1995.
- [166] H. Zhu, H. Zang, K. Zhu, and B. Mukherjee. A novel generic graph model for traffic grooming in heterogeneous WDM mesh network. *IEEE/ACM Transactions on Networking*, 11(2):285–299, April 2003.
- [167] K. Zhu and B. Mukherjee. Dynamic traffic grooming in optical WDM mesh networks with distributed control. In *Proc. SPIE APOC*, Beijing, China, November 2001.

- [168] K. Zhu and B. Mukherjee. On-line approaches for provisioning connections of different bandwidth granularities in WDM mesh networks. In *Proc. IEEE OFC*, Anaheim, CA, March 2002.
- [169] K. Zhu and B. Mukherjee. Traffic grooming in an optical WDM mesh network. *IEEE Journal on Selected Areas in Communications*, 20(1):122–133, January 2002.

Apêndice A

Análise Preliminar

A.1 Introdução

Neste apêndice são apresentadas algumas propriedades das soluções do **TGP**, além de outras relativas a estrutura do grafo utilizado na representação estendida da rede (bem como, na representação em camadas). Tais resultados foram obtidos a partir da análise preliminar da estrutura facial do poliedro associado ao conjunto de soluções do **TGP**, podendo ser encarados como subprodutos de tal procedimento. Na próxima seção examinam-se propriedades das soluções viáveis do **TGP**; enquanto que, na seção seguinte, apresentam-se os resultados relativos a estrutura do grafo $G = (N, A)$ utilizado na representação estendida da rede, bem como aqueles relativos ao grafo $G_\ell = (N_\ell, A_\ell)$ utilizado na representação em camadas da rede.

A.2 Sobre as Soluções

Seja \mathcal{S} o conjunto de soluções viáveis do **TGP**. Dessa forma, para a formulação apresentada na subseção 3.1.4, TGP_C , temos que:

$$\mathcal{S} = \{(w, f) \in \{0, 1\}^{|A_\ell|} \times \mathbb{Z}^{|A_\ell| \times |P|} \mid (w, f) \text{ satisfaz a (3.10b)–(3.10g)}\}.$$

O teorema da decomposição de fluxo [1] permite obter, para um problema envolvendo um só produto, um conjunto (não necessariamente único) de fluxos em caminhos e em ciclos a partir dos fluxos nos arcos. Tal teorema pode ser estendido da seguinte forma para um fluxo de vários produtos.

Proposição A.1. *Seja $(w, f) \in \mathcal{S}$, então o vetor f de fluxos nos arcos pode ser decomposto em um conjunto de caminhos \mathcal{C}_{cam} e um conjunto de ciclos $\mathcal{C}_{\text{ciclo}}$ (não necessariamente únicos), sendo que todos os caminhos com fluxo positivo se originam em um nó de oferta e terminam em um nó de demanda.*

Demonstração. Dado o vetor f , é possível se obter um conjunto de caminhos, $\mathcal{C}_{\text{cam}}^p$, e de ciclos, $\mathcal{C}_{\text{ciclo}}^p$, (não necessariamente únicos) para cada $t_p \in P$ aplicando o teorema da decomposição de fluxo [1]. Em seguida, basta fazer $\mathcal{C}_{\text{cam}} = \bigcup_{t_p \in P} \mathcal{C}_{\text{cam}}^p$ e $\mathcal{C}_{\text{ciclo}} = \bigcup_{t_p \in P} \mathcal{C}_{\text{ciclo}}^p$. \square

Caso se assuma que todos os custos são não-negativos, então não haverá nenhum ciclo cujo custo seja negativo (o custo de instalar as capacidades necessárias para transportar fluxo através do ciclo) e a seguinte proposição é verdadeira.

Proposição A.2. *Se para uma instância do TGP, todos os ciclos possuírem custo não-negativo e o conjunto de soluções não for vazio então existe um solução ótima $(\bar{w}, \bar{f}) \in \mathcal{S}$ para qual o conjunto de ciclos $\mathcal{C}_{\text{ciclo}}$ obtidos a partir de \bar{f} é vazio.*

Demonstração. Para qualquer solução $(w, f) \in \mathcal{S}$, o fluxo em um ciclo pode assumir qualquer valor não-negativo sem violar as restrições de conservação de fluxo (3.10b). Por outro lado, as restrições (3.10d) permitem que qualquer fluxo positivo em um ciclo seja reduzido a zero mantendo-se todas as restrições do modelo satisfeitas, gerando, assim, uma nova solução. Essa nova solução possui um custo igual ou menor que a anterior uma vez que o custo associado a instalação das capacidades necessárias para transportar fluxo através do ciclo é não negativo. Como o conjunto de soluções não é vazio, então existe solução para o problema. Considere, em particular, $(\hat{w}, \hat{f}) \in \mathcal{S}$ uma solução ótima em que o conjunto de ciclos não é vazio, logo pode-se aplicar uma linha de raciocínio análoga a anterior para mostrar que existe uma solução $(\bar{w}, \bar{f}) \in \mathcal{S}$, obtida a partir de (\hat{w}, \hat{f}) , para qual o conjunto de ciclos $\mathcal{C}_{\text{ciclo}}$ obtidos a partir de \bar{f} é vazio. \square

A proposição anterior garante a existência de uma solução ótima $(\bar{w}, \bar{f}) \in \mathcal{S}$ para qual o conjunto de ciclos $\mathcal{C}_{\text{ciclo}}$ obtidos a partir de \bar{f} é vazio, isto é, em que apenas caminhos simples são utilizados. Além disso, pode-se limitar o comprimento de tais caminhos utilizando-se da seguinte proposição.

Proposição A.3. *Seja $(w, f) \in \mathcal{S}$ e \mathcal{C}_{cam} o conjunto de caminhos obtidos a partir de f , então todo $\bar{c} \in \mathcal{C}_{\text{cam}}$ possui tamanho igual a $3v_{\bar{c}} - v_w - 3$, em que $v_{\bar{c}}$ representa o número de nós do caminho na rede original $H = (V, E)$ (utilizada para gerar a representação em camadas) e v_w , o número de nós do caminho na rede original que utilizam o roteamento por comprimento de onda (“wavelength routing”).*

Demonstração. O número de arcos entre os nós do caminho na rede original é $v_{\bar{c}} - 1$. Cada um deles será representado por um arco distinto na representação em camadas. Cada nó intermediário (diferente das extremidades) do caminho na rede original corresponderá a um ou dois arcos. Caso se utilize de roteamento por comprimento de onda um único arco da rede em camadas será utilizado no caminho. Sendo assim, tem-se v_w arcos correspondentes ao uso do roteamento por comprimento de ondas. Já para cada um dos $v_{\bar{c}} - v_w - 2$ nós intermediários, em que não se utiliza do roteamento por comprimento de onda, serão necessários 2 arcos na rede em camadas de modo a representar as operações de retirada (*drop*) e inserção (*add*), totalizando $2(v_{\bar{c}} - v_w - 2)$. Finalmente, cada nó extremo do caminho corresponderá apenas a um arco na rede em camadas, uma vez que na origem será representada apenas a operação de inserção (*add*), enquanto que no destino se realiza apenas a operação de retirada (*drop*). Sendo assim o número total de arcos de um caminho na representação em camadas correspondente a um caminho na rede original com $v_{\bar{c}}$ nós dos quais v_w se utilizam do roteamento por comprimento de onda é igual a $v_{\bar{c}} - 1 + v_w + 2(v_{\bar{c}} - v_w - 2) + 2 = 3v_{\bar{c}} - v_w - 3$. \square

Sendo assim, pode-se limitar o tamanho dos caminhos simples obtidos a partir de uma solução $(w, f) \in \mathcal{S}$ através do seguinte corolário.

Corolário A.4. *Seja $(w, f) \in \mathcal{S}$ e \mathcal{C}_{cam} o conjunto de caminhos obtidos a partir de f , então todo $\bar{c} \in \mathcal{C}_{\text{cam}}$ possui tamanho (em número de arcos) entre $2v_{\bar{c}} - 1$ e $3v_{\bar{c}} - 3$.*

Demonstração. Em um caminho com $v_{\bar{c}}$ nós na rede original, o número de nós do caminho na rede original que utilizam o roteamento por comprimento de onda, isto é, v_w pode variar de 0 a $v_{\bar{c}} - 2$. Substituindo-se esses valores na equação do tamanho de um caminho ($= 3v_{\bar{c}} - v_w - 3$), obtém-se os limites mencionados. \square

Por fim, deve-se destacar que o número de vértices de um caminho simples na rede original $H = (V, E)$ varia entre 2 e v , fazendo, portanto, que todo caminho $\bar{c} \in \mathcal{C}_{\text{cam}}$ possua tamanho entre 3 e $3v - 3$ arcos.

A.3 Sobre a Estrutura do Grafo

Nesta seção são apresentadas propriedades importantes sobre a estrutura do grafo $G = (N, A)$ associado à representação estendida da rede. Em especial, deseja-se explorar as propriedades e características relativas à conectividade do mesmo. Em seguida, procura-se estender tais resultados para o grafo $G_\ell = (N_\ell, A_\ell)$ utilizado na representação em camadas da rede.

Em primeiro lugar, cumpre destacar que o direcionamento dos arcos de $G = (N, A)$ será desprezado e, portanto, deve-se introduzir a seguir a noção de *grafo subjacente* a um grafo direcionado.

Definição A.5. *Grafo subjacente a um grafo direcionado $G = (N, A)$ é o grafo não direcionado $G_S = (N_S, A_S)$, em que $N_S = N$ e se o arco não direcionado $(i, j) \in A_S$ então $(i, j) \in A$ ou $(j, i) \in A$.*

Uma vez que se desconsidere as direções dos arcos da representação estendida $G = (N, A)$, pode-se demonstrar que seu grafo subjacente é biconexo em arcos, o que implica na existência de pelo menos dois caminhos (não direcionados) entre cada par de nós cujos arcos são todos distintos. A proposição seguinte formaliza tal resultado.

Proposição A.6. *O grafo subjacente à representação estendida $G = (N, A)$ é biconexo em arcos.*

Demonstração. Seja $G_S = (N_S, A_S)$ o grafo subjacente a $G = (N, A)$, logo $N_S = N$ e se $(i, j) \in A_S$ então $(i, j) \in A$ ou $(j, i) \in A$. Considere que o grafo não direcionado $H = (V, E)$ seja uma representação da topologia da rede original e que g_{i_v} representa o grau de $i_v \in V$, ou ainda, $\text{grau}(i_v) = g_{i_v}$. Seja $G'_S = (N'_S, A'_S)$ um subgrafo de $G_S = (N_S, A_S)$ em que $N'_S = N_S$ e $A'_S \subseteq A_S$ é formado pelos arcos entre nós óticos associados a elementos distintos da rede e pelos arcos entre os nós de *add-drop* e seus respectivos nós óticos. Por construção, o número de nós óticos (de entrada e de saída) associados a um nó de *add-drop* $i \in N$ (correspondente ao nó $i_v \in V$) na representação estendida é $2g_{i_v}$, logo $\text{grau}(i_S) = 2g_{i_v}, \forall i_S \in N'_S$ correspondente a um

nó de *add-drop* $i \in N$. Por outro lado, cada um dos nós óticos de entrada (saída) $j \in N$ associados ao nó de *add-drop* $i \in N$ (correspondente ao nó $i_v \in V$) possui grau de entrada (saída) igual a 1 e grau de saída (entrada) igual a g_{i_v} . Dessa forma, $\text{grau}(j_S) = g_{i_v} + 1, \forall j_S \in N_S$ correspondente a um nó ótico $j \in N$. Em particular, $\text{grau}(j_S) = 2, \forall j_S \in N'_S$ correspondente a um nó ótico $j \in N$, uma vez que, por construção, em $G'_S = (N'_S, A'_S)$ só existem 2 arcos incidentes a cada um dos nós óticos. Sendo assim, $\text{grau}(i_S)$ é par $\forall i_S \in N'_S$ e, conseqüentemente, G'_S possui um ciclo euleriano (não necessariamente simples) fazendo com que não existam pontes em G'_S , ou ainda, G'_S é biconexo em arcos. Como G'_S é um subgrafo de G_S , pode-se concluir que G_S é biconexo em arcos, uma vez que o acréscimo de arcos (necessário para transformar G'_S em G_S) não reduz a conectividade do grafo. \square

De modo a se demonstrar que o grafo subjacente à representação estendida $G = (N, A)$ também é biconexo em nós (implicando na existência de pelo menos dois caminhos totalmente disjuntos entre cada par de nós), é necessário antes se explorar propriedades de alguns subgrafos de $G = (N, A)$. Sendo assim, deve-se introduzir o subgrafo de $G = (N, A)$ induzido pelo conjunto de nós óticos (de entrada e de saída) associados ao elemento $i \in V$ da rede $H = (V, E)$, doravante denominado de $G^i = (N^i, A^i)$.

Definição A.7. *Seja $G^i = (N^i, A^i), \forall i \in V$, o subgrafo da representação estendida $G = (N, A)$ induzido pelo conjunto de nós óticos (de entrada e de saída) associados ao elemento $i \in V$ da rede $H = (V, E)$. Dessa forma, $N^i = \{j \in N \mid j \text{ é nó ótico associado a } i\}$; enquanto que $A^i = \{(i, j) \in A \mid i \in N^i, j \in N^i\}$.*

A proposição seguinte garante que o grafo subjacente a $G^i = (N^i, A^i)$ é hamiltoniano para todo $i \in V$, tal que $\text{grau}(i) \geq 3$.

Proposição A.8. *O grafo subjacente a $G^i = (N^i, A^i)$ é hamiltoniano para todo $i \in V$, tal que $\text{grau}(i) \geq 3$.*

Demonstração. Por construção, o número de nós óticos associados a um elemento da rede $i \in V$ é $2g_i$, em que g_i representa o grau de i em $H = (V, E)$. Além disso, só existem arcos saindo dos nós de entrada e chegando aos nós de saída. Sendo assim, $G^i = (N^i, A^i)$ é bipartite (N^i pode ser dividido em dois subconjunto disjuntos: um contendo apenas nós óticos de entrada e outro apenas nós óticos de saída) e sempre possui um número par de nós. Seja $G^i_S = (N^i_S, A^i_S)$ o grafo subjacente a $G^i = (N^i, A^i)$. Considere, ainda, que e_j e $s_j, j = 1, \dots, g_i$, representam os nós correspondentes aos nós óticos de entrada e de saída respectivamente. Caso $g_i \geq 3$ seja ímpar, o ciclo

$$\langle e_1, s_2, e_3, \dots, s_{g_i-1}, e_{g_i}, s_1, e_2, s_3, \dots, e_{g_i-1}, s_{g_i}, e_1 \rangle$$

existe em G^i_S e é hamiltoniano. Por outro lado, se $g_i \geq 3$ for par, o ciclo

$$\langle e_1, s_2, e_3, s_4, \dots, e_{g_i-1}, s_{g_i}, e_2, s_1, e_{g_i}, s_{g_i-1}, e_{g_i-2}, s_{g_i-3}, \dots, e_4, s_3, e_1 \rangle$$

existe em G^i_S e é hamiltoniano. Logo, para $g_i \geq 3$, o grafo subjacente a $G^i = (N^i, A^i)$ é hamiltoniano. \square

Uma consequência imediata da proposição anterior é a existência de caminhos disjuntos entre cada par de nós pertencentes ao grafo subjacente a $G^i = (N^i, A^i)$, $i \in V$, em que $\text{grau}(i) \geq 3$. Esse resultado é formalizado a seguir.

Corolário A.9. *Para todo par de nós $x, y \in N^i$, existem dois caminhos disjuntos no grafo subjacente a $G^i = (N^i, A^i)$, $i \in V$, tal que $\text{grau}(i) \geq 3$.*

Demonstração. Como o grafo subjacente a $G^i = (N^i, A^i)$ é hamiltoniano, então existe um ciclo (ver prova da Prop. A.8) no grafo subjacente que contém todo par de nós $x, y \in N^i$. Esse ciclo sempre poderá ser subdividido em um par de caminhos disjuntos entre x e y , o que demonstra a existência de pelo menos dois caminhos disjuntos entre cada de nós $x, y \in N^i$ no grafo subjacente a G^i . \square

O próximo passo em direção à prova da biconectividade em nós do grafo subjacente a $G = (N, A)$ é incluir o nó de *add-drop* nos resultados anteriores (Prop. A.8 e Cor. A.9). Para tanto, torna-se necessário definir $\bar{G}^i = (\bar{N}^i, \bar{A}^i)$, $\forall i \in V$, como o subgrafo de $G = (N, A)$ induzido pelo nó de *add-drop* juntamente com os nós óticos (de entrada e de saída) associados ao elemento $i \in V$ da rede $H = (V, E)$.

Definição A.10. *Seja $\bar{G}^i = (\bar{N}^i, \bar{A}^i)$, $\forall i \in V$, o subgrafo da representação estendida $G = (N, A)$ induzido pelo nó de *add-drop* juntamente com os nós óticos (de entrada e de saída) associados ao elemento $i \in V$ da rede $H = (V, E)$. Dessa forma, $\bar{N}^i = N^i \cup \{i_e\}$, em que i_e representa o nó de *add-drop* correspondente a $i \in V$ na representação estendida $G = (N, A)$; enquanto que $\bar{A}^i = \{(i, j) \in A \mid i \in \bar{N}^i, j \in \bar{N}^i\}$.*

De forma análoga a Prop. A.8, a proposição seguinte garante que o grafo subjacente a $\bar{G}^i = (\bar{N}^i, \bar{A}^i)$ é hamiltoniano para todo $i \in V$, tal que $\text{grau}(i) \geq 3$.

Proposição A.11. *O grafo subjacente a $\bar{G}^i = (\bar{N}^i, \bar{A}^i)$ é hamiltoniano para todo $i \in V$, tal que $\text{grau}(i) \geq 3$.*

Demonstração. Seja $\bar{G}_S^i = (\bar{N}_S^i, \bar{A}_S^i)$ o grafo subjacente a $\bar{G}^i = (\bar{N}^i, \bar{A}^i)$. Considere, ainda, que e_j e s_j , $j = 1, \dots, g_i$, representam os nós correspondentes aos nós óticos de entrada e de saída, respectivamente, e que i_e representa o nó correspondente ao nó de *add-drop*. Como existe um arco entre i_e e cada um dos nós correspondentes a nós óticos (de entrada e de saída), basta remover um dos arcos do ciclo hamiltoniano obtido na prova da Prop. A.8, substituindo-o por um caminho passando por i_e que conecte as extremidades do arco removido. Caso $g_i \geq 3$ seja ímpar, substituindo-se o último arco do ciclo apresentado na prova da Prop. A.8, obtém-se o ciclo

$$\langle e_1, s_2, e_3, \dots, s_{g_i-1}, e_{g_i}, s_1, e_2, s_3, \dots, e_{g_i-1}, s_{g_i}, i_e, e_1 \rangle$$

que existe em \bar{G}_S^i e é hamiltoniano. Por outro lado, se $g_i \geq 3$ for par, substituindo-se o último arco do ciclo apresentado na prova da Prop. A.8, obtém-se o ciclo

$$\langle e_1, s_2, e_3, s_4, \dots, e_{g_i-1}, s_{g_i}, e_2, s_1, e_{g_i}, s_{g_i-1}, e_{g_i-2}, s_{g_i-3}, \dots, e_4, s_3, i_e, e_1 \rangle$$

que existe em \bar{G}_S^i e é hamiltoniano. Logo, para $g_i \geq 3$, o grafo subjacente a $\bar{G}^i = (\bar{N}^i, \bar{A}^i)$ é hamiltoniano. \square

Além disso, de forma análoga ao Cor. A.9, uma consequência imediata da proposição anterior é a existência de caminhos disjuntos entre cada par de nós pertencentes ao grafo subjacente a $\bar{G}^i = (\bar{N}^i, \bar{A}^i)$, $i \in V$, em que $\text{grau}(i) \geq 3$. Esse resultado é formalizado no seguinte corolário.

Corolário A.12. *Para todo par de nós $x, y \in \bar{N}^i$, existem dois caminhos disjuntos no grafo subjacente a $\bar{G}^i = (\bar{N}^i, \bar{A}^i)$, $i \in V$, tal que $\text{grau}(i) \geq 3$.*

Demonstração. Como o grafo subjacente a $\bar{G}^i = (\bar{N}^i, \bar{A}^i)$ é hamiltoniano, então existe um ciclo (ver prova da Prop. A.11) no grafo subjacente que contém todo par de nós $x, y \in \bar{N}^i$. Esse ciclo sempre poderá ser subdividido em um par de caminhos disjuntos entre x e y , o que demonstra a existência de pelo menos dois caminhos disjuntos entre cada de nós $x, y \in \bar{N}^i$ no grafo subjacente a \bar{G}^i . \square

Em seguida, apresenta-se duas proposições cujas demonstrações são relativamente simples e, portanto, foram omitidas neste texto.

Proposição A.13. *A remoção de um nó ótico da representação estendida $G = (N, A)$ associado ao elemento $i \in V$ da rede $H = (V, E)$ não torna o grafo subjacente a ele desconexo, para todo $i \in V$, tal que $\text{grau}(i) < 3$.*

Proposição A.14. *A remoção de um nó de add-drop da representação estendida $G = (N, A)$ associado ao elemento $i \in V$ da rede $H = (V, E)$ não torna o grafo subjacente a ele desconexo, para todo $i \in V$, tal que $\text{grau}(i) < 3$.*

Finalmente, pode-se iniciar a demonstração relativa a biconectividade em nós do grafo subjacente à representação estendida $G = (N, A)$. Para tanto, serão utilizados dois lemas apresentados a seguir. O primeiro desses lemas demonstra que a remoção de um nó de *add-drop* da representação estendida $G = (N, A)$ não é suficiente para tornar o grafo subjacente a ele desconexo. A demonstração é dada a seguir.

Lema A.15. *A remoção de um nó de add-drop da representação estendida $G = (N, A)$ não torna o grafo subjacente a ele desconexo.*

Demonstração. Seja $G_S = (N_S, A_S)$ o grafo subjacente a $G = (N, A)$, logo $N_S = N$ e se $(i, j) \in A_S$ então $(i, j) \in A$ ou $(j, i) \in A$. Suponha que $G'_S = (N'_S, A'_S)$ é desconexo, em que $N'_S = N_S \setminus \{i_S\}$, tal que i_S é um nó correspondente a um nó de *add-drop* $i \in N$. Assim, $A'_S = A_S \setminus \{(i_S, j) \mid j \in N'_S\}$. Como G'_S é desconexo, existe um par de nós $u, v \in N'_S$, tal que não há caminho entre eles. Caso u e v sejam nós correspondentes a nós óticos associados a i em $G(N, A)$, sempre existirá um caminho *alternativo* entre eles (que não passa por i_S), uma vez que G'_S é hamiltoniano. Caso u e v sejam nós (óticos ou não) correspondentes a outros elementos da rede, então a remoção de i_S só tornará o grafo desconexo caso exista apenas um único caminho entre u e v em G_S . Seja \bar{c} o caminho em G_S entre u e v passando por i_S . Como i_S é um nó correspondente a um nó de *add-drop*, então \bar{c} é da forma $\langle u, \dots, t, i_S, w, \dots, v \rangle$, ou melhor, existem t e w pertencentes ao caminho que são nós correspondentes a nós óticos associados ao mesmo elemento de rede que i_S . Como argumentado acima, sempre existirá um caminho *alternativo* (que não passa por i_S) entre t e w . Daí, conclui-se que sempre existirá um caminho entre u e v , mesmo após a remoção de i_S , resultando em uma contradição. Logo $G'_S = (N'_S, A'_S)$ é conexo. \square

O segundo lema, por sua vez, garante que a remoção de um nó ótico da representação estendida $G = (N, A)$ também não é suficiente para tornar o grafo subjacente a ele desconexo, como pode ser visto a seguir.

Lema A.16. *A remoção de um nó ótico da representação estendida $G = (N, A)$ não torna o grafo subjacente a ele desconexo.*

Demonstração. Seja $G_S = (N_S, A_S)$ o grafo subjacente a $G = (N, A)$, logo $N_S = N$ e se $(i, j) \in A_S$ então $(i, j) \in A$ ou $(j, i) \in A$. Suponha que $G'_S = (N'_S, A'_S)$ é desconexo, em que $N'_S = N_S \setminus \{i_S\}$, tal que i_S é um nó correspondente a um nó ótico no elemento de rede $i \in V$. Como G'_S é desconexo, existe um par de nós $u, v \in N'_S$, tal que não há caminho entre eles em G'_S . Seja \bar{c} o caminho em G_S entre u e v passando por i_S , tal que $\bar{c} = \langle u, \dots, t, i_S, w, \dots, v \rangle$.

Caso t e w sejam nós correspondentes a nós óticos do mesmo elemento de rede, então existe um caminho alternativo (que não passa por i_S) entre t e w (ver Cor. A.9). Logo a remoção de i_S não elimina todos os caminhos entre u e v .

Caso t (ou w) seja um nó correspondente ao nó de *add-drop* do mesmo elemento de rede que i_S e w (ou t) seja nó correspondente a um nó ótico do mesmo elemento de rede que i_S , por razões semelhantes (ver Cor. A.12) pode-se concluir que existe um caminho *alternativo* entre u e v .

Caso t (ou w) seja um nó correspondente a um nó ótico de um elemento de rede diferente daquele associado a i_S e w (ou t) seja nó correspondente a um nó ótico do mesmo elemento de rede que i_S , também existirá um caminho *alternativo* entre u e v que não passa por i_S . Sem perda de generalidade, suponha que t seja um nó correspondente a um nó ótico de saída (logo i_S corresponde a um nó de entrada) associado ao elemento de rede $l \in V$. Como o grafo subjacente a G^l é hamiltoniano, então existe um caminho entre t e x , em que x é o nó correspondente ao nó ótico de entrada associado a mesma conexão que t no elemento de rede l . Seja y o nó correspondente ao nó ótico de saída pertencente ao mesmo elemento de rede que i_S e associado a mesma conexão que i_S . Como y e w são nós correspondentes a nós óticos do mesmo elemento de rede, então $y \neq w$, pois, por construção, não existe arco entre y e i_S . Contudo, como G^i é hamiltoniano, existe um caminho entre y e w que não passa por i_S . Daí, pode-se concluir que existe um caminho *alternativo* entre u e v , ou melhor, $\langle u, \dots, t, \dots, x, y, \dots, w, \dots, v \rangle$ que não passa por i_S .

Caso t (ou w) seja um nó correspondente a um nó ótico de um elemento de rede diferente daquele associado a i_S e w (ou t) seja nó correspondente ao nó de *add-drop* do mesmo elemento de rede que i_S , também existirá um caminho *alternativo* entre u e v que não passa por i_S . Sem perda de generalidade, suponha que t seja um nó correspondente a um nó ótico de saída (logo i_S corresponde a um nó de entrada) associado ao elemento de rede $l \in V$. Como o grafo subjacente a G^l é hamiltoniano, então existe um caminho entre t e x , em que x é o nó correspondente ao nó ótico de entrada associado a mesma conexão que t no elemento de rede l . Seja y o nó correspondente ao nó ótico de saída pertencente ao mesmo elemento de rede que i_S e associado a mesma conexão que i_S . Como y é um nó correspondente a um nó ótico do mesmo elemento de rede que w , então, por construção, existe um arco entre y e w . Daí, pode-se concluir que existe um caminho *alternativo* entre u e v , ou melhor, $\langle u, \dots, t, \dots, x, y, w, \dots, v \rangle$ que não passa por i_S .

Finalmente, conclui-se que sempre existirá um caminho entre u e v , mesmo após a remoção de i_S , resultando em uma contradição. Logo $G'_S = (N'_S, A'_S)$ é conexo. \square

Por fim, utilizando-se os Lemas A.15 e A.16, pode-se demonstrar que o grafo subjacente à representação estendida $G = (N, A)$ é biconexo em nós. Tal resultado é formalizado na seguinte proposição.

Proposição A.17. *O grafo subjacente à representação estendida $G = (N, A)$ é biconexo em nós.*

Demonstração. Seja $G_S = (N_S, A_S)$ o grafo subjacente a $G = (N, A)$. Para se mostrar que G_S é biconexo em nós basta mostrar que não existem articulações em G_S , isto é, não existe $i \in N_S$ cuja remoção torne G_S desconexo. Os Lemas A.15 e A.16 garantem que a remoção de um nó (ótico ou de *add-drop*) não torna o grafo subjacente desconexo, logo pode-se concluir que não existem articulações em G_S . Portanto, G_S é biconexo em nós. \square

Conforme mencionado no início desta seção, procurou-se estender os resultados obtidos para o grafo $G_\ell = (N_\ell, A_\ell)$ utilizado na representação em camadas da rede. Neste caso, as proposições e suas respectivas demonstrações são muito similares as que foram apresentadas anteriormente (e foram omitidas neste texto). Contudo, vale estabelecer (sem apresentar sua demonstração) o seguinte resultado.

Proposição A.18. *O grafo subjacente à representação em camadas $G_\ell = (N_\ell, A_\ell)$ é biconexo.*

Além disso, caso o grafo correspondente à rede original, $H = (V, E)$, seja biconexo a seguinte proposição também é verdadeira.

Proposição A.19. *Se o grafo correspondente à rede original $H = (V, E)$ for biconexo, então no grafo da representação em camadas $G_\ell = (N_\ell, A_\ell)$ existem 2 caminhos disjuntos para cada par de nós de *add-drop*.*

Cumpramos ressaltar que tal condição (biconectividade da rede original) é comumente atendida pela maioria dos projetos atuais de redes óticas, de forma a se obter uma topologia física que suporte eventuais falhas, como a queda de um enlace físico ou a parada de um dos elementos da rede.

A.4 Consideração Finais

Neste apêndice procurou-se apresentar algumas propriedades das soluções do **TGP**, além de outras relativas a estrutura do grafo utilizado na representação estendida da rede (bem como, na representação em camadas). Tais resultados foram obtidos durante o processo de análise da estrutura facial do poliedro associado ao conjunto de soluções do **TGP**. Mais ainda, alguns deles são essenciais a esse estudo e à demonstração de alguns resultados apresentados no capítulo 6.