

Heurística Lagrangeana para o Problema de Traffic Grooming em Redes Óticas WDM

Zenilton Kleber G. Patrocínio Jr.¹, Geraldo Robson Mateus¹

¹Departamento de Ciência da Computação – Universidade Federal de Minas Gerais
Caixa Postal 702 – 30123-970 Belo Horizonte, MG

zenilton@dcc.ufmg.br, mateus@dcc.ufmg.br

Abstract. *In a wavelength division multiplexing (WDM) optical network, each wavelength can carry several lower-rate traffic streams. The problem of combining low speed traffic streams into high speed ones in order to fully explore available capacity has been called Traffic Grooming Problem (TGP). In this paper, TGP in WDM optical networks regardless of underlying physical topology is investigated. The problem is formulated as an integer linear program and two versions of a Lagrangian-based heuristic are presented. This heuristic divides the original problem into simple subproblems which are solved independently. Computational experiments show that our Lagrangian-based heuristic can obtain better results than other heuristics presented in the literature.*

Resumo. *Em uma rede ótica utilizando multiplexação por divisão de comprimento de onda (WDM – Wavelength Division Multiplexing), cada comprimento de onda pode transportar vários tráfegos de baixa velocidade. O problema de se combinar os tráfegos de baixa velocidade de modo a se obter uma melhor ocupação da capacidade de transporte disponível é conhecido na literatura como Traffic Grooming Problem (TGP). Neste artigo, investiga-se o TGP em redes óticas WDM independentemente da topologia física subjacente. Para tanto, apresenta-se uma formulação do problema através de um programa linear inteiro (PLI), além de duas versões de uma heurística lagrangeana (HL) para sua resolução. Essa heurística divide o problema original em subproblemas mais simples que são resolvidos de forma independente. Os experimentos computacionais indicam que a HL é capaz de obter resultados melhores que outras heurísticas encontradas na literatura.*

1. Introdução

A multiplexação por divisão de comprimento de onda (WDM – “wavelength division multiplexing”) tem se estabelecido como uma técnica extremamente útil e poderosa a ser utilizada no planejamento, projeto e implementação das atuais redes óticas, na medida que seu uso permite um aumento da capacidade de transmissão através das fibras óticas. Atrasos normalmente devidos às limitações de velocidade dos componentes eletrônicos localizados nos nós da rede passariam a ser superados através do uso de diferentes comprimentos de onda para o estabelecimento de “canais independentes” no mesmo meio físico, isto é, na mesma fibra [Dutta and Rouskas 2000]. Além disso, em virtude da limitação do número de comprimentos de onda disponíveis na tecnologia atual, não é viável se atribuir um comprimento de onda exclusivo para cada tráfego a ser transportado. Dessa forma,

de modo a se obter uma solução viável e mais barata para o problema, deve-se procurar combinar vários tráfegos de baixa velocidade de modo a se obter uma melhor ocupação da capacidade de transporte disponível, por exemplo através da multiplexação por divisão de tempo (TDM – “*time division multiplexing*”). Além do mais, esse procedimento deve ser realizado de maneira a minimizar a necessidade e uso de conversores ótico-eletrônicos (O–E–O) nos elementos intermediários do trajeto que liga a origem de um dado tráfego ao seu destino.

Esse problema de se combinar os tráfegos de baixa velocidade em tráfegos de alta velocidade, bem como estabelecer seu roteamento, tem sido designado de *Traffic Grooming Problem (TGP)* na literatura e, recentemente, tem sido alvo de estudo de vários autores, especialmente para topologias em anel [Wang et al. 2001, Dutta and Rouskas 2002]. A maioria desses trabalhos explora o uso de técnicas heurísticas na resolução desse problema. Mais recentemente, alguns trabalhos passaram a lidar com topologias irregulares, como, por exemplo, [Zhu and Mukherjee 2002]. Vale também mencionar que o problema de projeto de uma topologia lógica (ou virtual) é NP-difícil mesmo para o caso em que a topologia física é dada por um simples anel [Dutta and Rouskas 2000].

Neste artigo, uma heurística baseada na relaxação lagrangeana é apresentada e aplicada ao **TGP** em redes óticas **WDM** independentemente da topologia física subjacente. Para tanto, **TGP** foi formulado através de um programa linear inteiro (**PLI**) utilizando-se uma representação da topologia de rede através de um grafo em camadas. A relaxação lagrangeana foi usada para se gerar limites inferiores a serem utilizados em um procedimento de busca. Uma heurística que se utiliza de informações dos limites inferiores para obtenção de limites superiores de “boa qualidade” também foi desenvolvida. Os resultados obtidos foram superiores se comparados aos de outra abordagem encontrada na literatura.

O restante deste artigo está organizado da seguinte forma. A seção 2 descreve a construção de uma representação em camadas para a topologia de rede. Em seguida, apresenta-se a formulação matemática do **TGP** na seção 3. A heurística lagrangeana é descrita na seção 4, enquanto que os resultados dos experimentos computacionais são descritos na seção 5. Finalmente, a seção 6 sumariza os principais resultados obtidos neste artigo.

2. Representação através de um Grafo em Camadas

Apesar de um grafo não direcionado ser suficiente para representar as informações básicas sobre a topologia de uma rede ótica (isto é, a disposição dos componentes óticos e de suas interconexões), ele não permite a representação dos componentes internos (multiplexadores e comutadores) dos elementos da rede. Tal representação se torna necessária para se modelar o **TGP**.

Para tanto, cada nó da rede original será subdividido em vários outros, sendo representado na realidade por um conjunto de nós e arcos (ver Fig. 1). 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

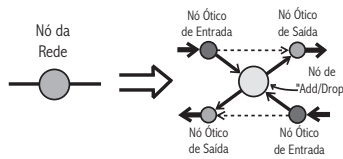


Figura 1. Subdivisão de um Nó

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 1 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-eleto-ótica). Na Figura 1 estas conexões estão representadas através de linhas pontilhadas entre os nós óticos de entrada e saída.

Baseado nessa representação estendida de cada elemento da rede, uma representação da topologia de rede através de um grafo em camadas é construída da seguinte forma. Cada *lambda* (ou comprimento de onda) será representado por uma camada distinta. Os nós óticos da representação estendida serão duplicados em cada camada, bem como as conexões (arcos) entre eles. Já cada um dos nós de *add-drop*, por sua vez, será mantido inalterado sem contudo estar associado a nenhuma camada em particular. As conexões entre estes e as duplicatas dos nós óticos serão estabelecidas da seguinte forma: para cada arco saindo do nó de *add-drop* i e entrando no nó ótico j da representação estendida será 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 será 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 rede com apenas dois nós exibida na Figura 2(a), obtém-se, após a adição dos nós óticos e de suas conexões, conforme descrito anteriormente, a representação estendida mostrada na Figura 2(b). Caso o número total de *lambdas* (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 2(c) exibe a representação em camadas correspondente a representação estendida da Figura 2(b) considerando um disponibilidade de 2 *lambdas*. Em tal figura, os nós óticos pertencentes a mesma camada, isto é, que se utilizam do mesmo *lambda*, encontram-se agrupados através de um retângulo pontilhado.

3. Formulação Matemática do TGP

A seguinte notação é utilizada na formulação matemática do **TGP** para uma rede ótica cuja topologia é representada através de um grafo em camadas $G = (N, A)$:

- N conjunto de nós da representação em camadas, tal que $N = N^e \cup N^o$, em que
- N^e representa o conjunto de nós de *add-drop* e N^o , o conjunto de nós óticos;

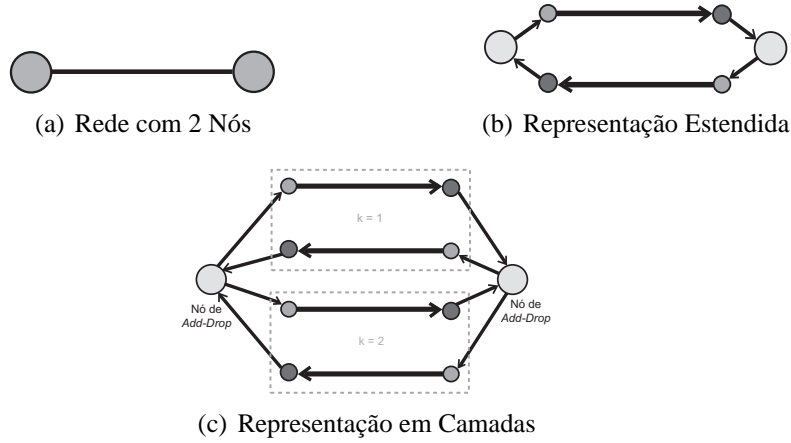


Figura 2. Exemplo de Representação em Camadas

- A conjunto de arcos da representação em camadas;
 P conjunto de requisições de tráfego, $P = \{t_p = (s_p, d_p, f_p, m_p)\}$, n_p representa o número total de requisições;
 s_p origem da requisição de tráfego $t_p \in P$;
 d_p destino da requisição de tráfego $t_p \in P$;
 f_p demanda de tráfego da requisição $t_p \in P$;
 m_p quantidade máxima de tráfego da requisição $t_p \in P$ que pode ser transportada através de um λ ;
 c_ℓ custo de alocação do arco $\ell \in A$;
 Δ capacidade total de um comprimento de onda, $\Delta = \text{mmc}^1\{m_p \mid \forall t_p \in P\}$;
 δ_p contribuição (ou ocupação) correspondente a uma unidade de tráfego transportada em um λ , $\delta_p = \Delta/m_p, \forall t_p \in P$.

Em verdade, P é um multiconjunto, isto é, P pode conter elementos iguais. Além disso, as seguintes variáveis serão utilizadas na formulação condensada do **TGP**:

- f_ℓ^p quantidade de tráfego da requisição $t_p \in P$ no arco $\ell \in A$;
 w_ℓ indica o uso do arco $\ell \in A$ no transporte de algum tráfego, isto é, indica se o arco pertence ou não à topologia virtual.

Por fim, $A^+(i)$ é utilizado para representar o conjunto de todos arcos que saem de um nó $i \in N$, isto é, $A^+(i) = \{\ell \in A \mid \ell = (i, j), j \in N\}, \forall i \in N$, enquanto que $A^-(i)$ representa o conjunto de todos os arcos que chegam em um nó $i \in N$, isto é, $A^-(i) = \{\ell \in A \mid \ell = (j, i), j \in N\}, \forall i \in N$.

Dessa forma, o **TGP** pode ser formulado através de um **PLI** dado por:

$$(TGP) \quad \min \sum_{\ell \in A} c_\ell w_\ell \quad (1a)$$

sujeito a:

$$\sum_{\ell \in A^+(i)} f_\ell^p - \sum_{\ell \in A^-(i)} f_\ell^p = b_i^p, \quad \forall t_p \in P, \forall i \in N \quad (1b)$$

$$\sum_{\ell \in A^+(i)} w_\ell - \sum_{\ell \in A^-(i)} w_\ell = 0, \quad \forall i \in N^o \quad (1c)$$

¹Mínimo múltiplo comum.

$$\sum_{t_p \in P} \delta_p f_\ell^p \leq \Delta w_\ell, \forall \ell \in A \quad (1d)$$

$$f_\ell^p \geq 0, \forall t_p \in P, \forall \ell \in A \quad (1e)$$

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

$$f_\ell^p \text{ inteiro}, \forall t_p \in P, \forall \ell \in A \quad (1g)$$

em que b_i^p é dado por:

$$b_i^p = \begin{cases} f^p, & \text{se } i = s_p \\ -f^p, & \text{se } i = d_p \\ 0, & \text{se } i \neq s_p \neq d_p \end{cases}, \forall t_p \in P, \forall i \in N.$$

A função objetivo dada por (1a) procura minimizar o custo total de utilização/alocação dos arcos da representação em camadas. As restrições (1b) garantem a conservação de fluxo para todos os nós da representação em camadas. Já as restrições (1c) 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, portanto, responsáveis pela continuidade dos *lightpaths*. As restrições (1d), por sua vez, impõem um limite sobre o volume total de tráfego transportado através de um arco $\ell \in A$, além de estabelecer uma ligação entre as variáveis de fluxo (f_ℓ^p) e de decisão (w_ℓ). Finalmente, as restrições (1e), (1f) e (1g) 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.

4. Heurística Lagrangeana

Uma heurística lagrangeana (**HL**) pode ser vista como uma restrição da implementação de um algoritmo de *branch-and-bound* completo, em que se explora apenas o nó inicial (correspondente ao problema original). Nesse procedimento iterativo, procura-se produzir uma sequência de limitantes inferiores e superiores para a solução do problema de modo que, durante todo o processo, se possa avaliar a qualidade das soluções viáveis encontradas. Sendo assim, o **TGP** foi relaxado (via relaxação lagrangeana) para o cálculo de limites inferiores conforme descrito adiante e utilizou-se o método de subgradientes para implementação de um procedimento de busca (maximização dos limites inferiores). Durante tal busca, uma heurística “gulosa” baseada no cálculo de caminhos mínimos foi utilizada para a obtenção de novas soluções viáveis (limites superiores) para o problema.

O Algoritmo 1 descreve o funcionamento da **HL** cujos detalhes de implementação são dados a seguir.

4.1. Limite Inferior

De modo a se realizar a relaxação lagrangeana, um vetor de multiplicadores de Lagrange π , tal que $\pi_i^p \in R, \forall t_p \in P, \forall i \in N$, é associado às restrições de conservação de fluxo dos nós – restrições (1b). Já as restrições que limitam a capacidade dos arcos – restrições (1d), são associadas a um vetor de multiplicadores de Lagrange λ , tal que $\lambda_\ell \in R, \forall \ell \in A$.

4.1.1. Relaxação #1

No primeiro limite inferior, L_1^{TGP} , são relaxadas apenas as restrições (1d). Isto resulta na eliminação do relacionamento entre as variáveis de fluxo e de decisão. Mais ainda,

Algoritmo 1 Heurística Lagrangeana

procedure HL

$\hat{A}_0 \leftarrow A$

$\hat{L} \leftarrow \text{CALCULALIMITESUPERIOR}(\hat{A}_0)$

▷ Determina se existe solução

if $\hat{L} = +\infty$ **then**

PARE

▷ Problema é inviável

end if

$\hat{L}_{\min} \leftarrow \hat{L}$

▷ Inicializa melhor limite superior

$\text{MultiLagrange} \leftarrow \text{INICIALIZAMULTLAGRANGE}()$

▷ Inicializa multiplicadores

$L \leftarrow \text{CALCULALIMITEINFERIOR}(\text{MultiLagrange})$

▷ Inicializa melhor limite inferior

$L_{\max} \leftarrow L$

$\hat{A}_0 \leftarrow \{\ell \in A \mid \bar{w}_\ell = 1\}$

$\hat{L} \leftarrow \text{CALCULALIMITESUPERIOR}(\hat{A}_0)$

▷ Procura novo limite superior

$\hat{L}_{\min} \leftarrow \min\{\hat{L}_{\min}, \hat{L}\}$

$\text{Gap} \leftarrow (\hat{L}_{\min} - L_{\max})/\hat{L}_{\min}$

▷ Calcula *gap de dualidade*

$\text{Iter} \leftarrow 1$

$\text{Non_Increasing} \leftarrow 0$

$\sigma \leftarrow 1,1$

while ($\text{Gap} > \epsilon_1$) and ($\sigma > \epsilon_2$) and ($\text{Iter} < \text{Max_Iter}$) **do**

if $\text{Non_Increasing} > \text{Max_Non_Increasing}$ **then**

$\sigma \leftarrow \sigma/2$

$\text{Non_Increasing} \leftarrow 0$

end if

$\xi \leftarrow \text{CALCULASUBGRADIENTES}()$

if $\|\xi\| < \epsilon_3$ **then**

▷ Verifica se subgradiente é nulo

PARE

end if

$t \leftarrow \text{CALCULATAMANHOPASSO}(\sigma, \xi, L, \hat{L}_{\min})$

$\text{MultiLagrange} \leftarrow \text{ATUALIZAMULTLAGRANGE}(\text{MultiLagrange}, t, \xi)$

$L \leftarrow \text{CALCULALIMITEINFERIOR}(\text{MultiLagrange})$

if $L > L_{\max}$ **then**

▷ Verifica se novo limite inferior é melhor

$L_{\max} \leftarrow L$

▷ Atualiza melhor limite inferior

$\sigma \leftarrow 1,1$

$\text{Non_Increasing} \leftarrow 0$

end if

$\hat{A}_0 \leftarrow \{\ell \in A \mid \bar{w}_\ell = 1\}$

$\hat{L} \leftarrow \text{CALCULALIMITESUPERIOR}(\hat{A}_0)$

▷ Procura novo limite superior

$\hat{L}_{\min} \leftarrow \min\{\hat{L}_{\min}, \hat{L}\}$

$\text{Gap} \leftarrow (\hat{L}_{\min} - L_{\max})/\hat{L}_{\min}$

▷ Recalcula *gap de dualidade*

$\text{Iter} \leftarrow \text{Iter} + 1$

end while

end procedure

as variáveis de fluxo também deixam de ser interdependentes pois essas restrições são responsáveis pelo limite sobre o volume total de tráfego transportado através de um arco.

Dessa forma, o primeiro limite inferior, L_1^{TGP} , pode ser formulado da seguinte maneira:

$$L_1^{\text{TGP}}(\lambda) = \min \sum_{\ell \in A} c_\ell w_\ell + \sum_{\ell \in A} \lambda_\ell \left(\sum_{t_p \in P} \delta_p f_\ell^p - \Delta w_\ell \right)$$

sujeito a: (1b), (1c), (1e), (1f) e (1g).

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

$$L_1^{\text{TGP}}(\lambda) = L_{1W}^{\text{TGP}}(\lambda) + L_{1F}^{\text{TGP}}(\lambda) \quad (2)$$

em que:

$$L_{1W}^{\text{TGP}}(\lambda) = \min \sum_{\ell \in A} (c_\ell - \Delta \lambda_\ell) w_\ell \quad L_{1F}^{\text{TGP}}(\lambda) = \min \sum_{t_p \in P} \sum_{\ell \in A} \delta_p \lambda_\ell f_\ell^p$$

sujeito a: (1c) e (1f).

sujeito a: (1b), (1e) e (1g).

O primeiro subproblema (determinação da topologia virtual), L_{1W}^{TGP} , deve ser solucionado através da resolução de um problema de fluxo de custo mínimo. Antes porém, para se obter uma solução relaxada mais próxima da efetiva solução do problema, deve-se agregar as restrições (3a) e (3b) – inequações válidas – à formulação do subproblema L_{1W}^{TGP} dada anteriormente.

$$\sum_{\ell \in A^+(i)} w_\ell \geq \beta_i, \forall i \in N^e \quad (3a)$$

$$\sum_{\ell \in A^-(i)} w_\ell \geq \gamma_i, \forall i \in N^e \quad (3b)$$

em que β_i e $\gamma_i, \forall i \in N^e$, 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:

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

Como mencionado anteriormente, o subproblema $L_{1W}^{\text{TGP}}(\lambda)$ pode ser solucionado através da resolução de um problema de fluxo de custo mínimo. Nessa abordagem as facilidades instaladas nos arcos, isto é, os *lambdas* serão tratados como um único tipo de produto que deve fluir entre os nós de oferta e de demanda através de caminhos disjuntos, de modo a prevenir o *bloqueio*.

Seja \bar{w}_ℓ os valores das variáveis de decisão obtidos a partir da resolução do problema de fluxo de custo mínimo. Desse modo o valor da solução do primeiro subproblema, L_{1W}^{TGP} , é dado por:

$$L_{1W}^{\text{TGP}}(\lambda) = \sum_{\ell \in A} (c_\ell - \Delta \lambda_\ell) \bar{w}_\ell. \quad (4)$$

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

$$L_{\text{IF}}^{\text{TGP}}(\lambda) = \sum_{t_p \in P} \delta_p L_{\text{IF}_p}^{\text{TGP}}(\lambda) \quad (5)$$

em que para cada requisição de tráfego $t_p \in P$:

$$L_{\text{IF}_p}^{\text{TGP}}(\lambda) = \min \sum_{\ell \in A} \lambda_\ell f_\ell^p \quad (6a)$$

sujeito a:

$$\sum_{\ell \in A^+(i)} f_\ell^p - \sum_{\ell \in A^-(i)} f_\ell^p = b_i^p, \forall i \in N \quad (6b)$$

$$f_\ell^p \geq 0, \forall \ell \in A \quad (6c)$$

$$f_\ell^p \text{ inteiro}, \forall \ell \in A \quad (6d)$$

Cada um dos subproblemas de roteamento $L_{\text{IF}_p}^{\text{TGP}}(\lambda)$ 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 requisição. Define-se 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 cada requisição $t_p \in P$ através do caminho mínimo C^p . Seja \bar{f}_ℓ^p os valores obtidos, dessa maneira, para as variáveis de fluxo. Desse modo o valor da solução do segundo subproblema, $L_{\text{IF}}^{\text{TGP}}$, é dado por:

$$L_{\text{IF}}^{\text{TGP}}(\lambda) = \sum_{t_p \in P} \delta_p \sum_{\ell \in A} \lambda_\ell \bar{f}_\ell^p = \sum_{t_p \in P} \delta_p \sum_{\ell \in C^p} (\lambda_\ell f_p) = \sum_{t_p \in P} \delta_p f_p \left(\sum_{\ell \in C^p} \lambda_\ell \right). \quad (7)$$

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

$$L_1^{\text{TGP}}(\lambda) = L_{\text{1W}}^{\text{TGP}}(\lambda) + L_{\text{IF}}^{\text{TGP}}(\lambda) = \sum_{\ell \in A} (c_\ell - \Delta \lambda_\ell) \bar{w}_\ell + \sum_{t_p \in P} \delta_p f_p \left(\sum_{\ell \in C^p} \lambda_\ell \right). \quad (8)$$

4.1.2. Relaxação #2

No segundo limite inferior, L_2^{TGP} , são relaxadas as restrições (1d), juntamente com as restrições de conservação de fluxo (1b), sendo formulado da seguinte maneira:

$$L_2^{\text{TGP}}(\pi, \lambda) = \min \sum_{\ell \in A} c_\ell w_\ell + \sum_{t_p \in P} \sum_{i \in N} \pi_i^p \left(\sum_{\ell \in A^+(i)} f_\ell^p - \sum_{\ell \in A^-(i)} f_\ell^p - b_i^p \right) + \sum_{\ell \in A} \lambda_\ell \left(\sum_{t_p \in P} \delta_p f_\ell^p - \Delta w_\ell \right)$$

sujeito a: (1c), (1e), (1f) e (1g).

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

$$L_2^{\text{TGP}}(\pi, \lambda) = L_{\text{2W}}^{\text{TGP}}(\lambda) + L_{\text{2F}}^{\text{TGP}}(\pi, \lambda) + \sum_{t_p \in P} (\pi_{d_p}^p - \pi_{s_p}^p) f_p \quad (9)$$

em que:

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

$$L_{2F}^{\text{TGP}}(\pi, \lambda) = \sum_{t_p \in P} \sum_{\ell \in A} \tilde{c}_\ell^p f_\ell^p$$

sujeito a: (1c) e (1f).

sujeito a: (1e) e (1g).

em que $\tilde{c}_\ell^p = \delta_p \lambda_\ell + \pi_i^p - \pi_j^p, \forall \ell = (i, j) \in A, \forall t_p \in P$.

O primeiro subproblema (determinação da topologia virtual), L_{2W}^{TGP} , é equivalente ao primeiro subproblema da relaxação anterior (L_{1W}^{TGP}) e deve ser solucionado da mesma forma, isto é, através do acréscimo das inequações válidas (3a) e (3b) e resolução do problema de fluxo de custo mínimo associado.

Já o segundo subproblema (roteamento de tráfego), L_{2F}^{TGP} , pode ser resolvido por inspeção. Para cada requisição $t_p \in P$, basta verificar se o custo modificado do arco $\ell \in A$ é negativo, isto é, $\tilde{c}_\ell^p < 0$. Neste caso, deve-se rotear m_p unidades de tráfego da requisição t_p através do arco ℓ , caso contrário o fluxo é nulo.

Seja $(\bar{f}_\ell^p, \bar{w}_\ell)$ a solução obtida da forma descrita acima e A^p o conjunto de arcos $\ell \in A$ para os quais o custo modificado \tilde{c}_ℓ^p associado à requisição $t_p \in P$ é negativo, isto é, $A^p = \{\ell \in A \mid \tilde{c}_\ell^p < 0\}, \forall t_p \in P$. Sendo assim, o valor do segundo limite inferior, L_2^{TGP} , é dado por:

$$L_2^{\text{TGP}}(\pi, \lambda) = \sum_{\ell \in A} (c_\ell - \Delta \lambda_\ell) \bar{w}_\ell + \sum_{t_p \in P} m_p \left(\sum_{\ell \in A^p} \tilde{c}_\ell^p \right) + \sum_{t_p \in P} (\pi_{d_p}^p - \pi_{s_p}^p) f_p. \quad (10)$$

4.2. Limite Superior

De modo a se implementar um procedimento de busca 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 arcos da representação em camadas selecionados durante o cálculo do limite inferior como base para obtenção de uma solução viável.

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

A primeira parte da heurística roteia inicialmente cada requisição através do caminho mais curto entre sua origem e seu destino, utilizando-se apenas dos arcos pertencentes a \hat{A}_0 . Tal roteamento é realizado em ordem decrescente de ocupação da capacidade de transporte, isto é, a requisição cuja razão f^p/m_p for maior é roteada primeiro, sendo seguido por aquela 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 de uma requisição, um novo caminho é calculado e utilizado. Essa sequência de operações é repetida até que todo o tráfego da requisição tenha sido roteado ou que não se encontre um caminho de capacidade disponível entre a origem e o destino da requisição. Nesse último caso, considera-se que o conjunto de arcos selecionados 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 requisição, 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 (1c). 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 (1c). Para tanto, basta observar que o número de arcos alocados 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 buscam 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} o conjunto de arcos selecionados, isto é, $\hat{A} = \{\ell \in A \mid \hat{w}_\ell = 1\}$. Dessa forma, o valor do limite superior, \hat{L} , é dado por:

$$\hat{L} = \sum_{\ell \in \hat{A}} c_\ell. \quad (11)$$

4.3. Método de Subgradiente e Procedimento de Busca

Para cada conjunto de multiplicadores de Lagrange, obtém-se um limite inferior para o TGP, dado por $L_1^{\text{TGP}}(\lambda)$ ou $L_2^{\text{TGP}}(\pi, \lambda)$. 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:

$$(P_1^D) \quad \max_{\lambda \geq 0} L_1^{\text{TGP}}(\lambda), \quad (12)$$

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

$$(P_2^D) \quad \max_{\substack{\pi \in R \\ \lambda \geq 0}} L_2^{\text{TGP}}(\pi, \lambda). \quad (13)$$

O método de subgradientes é um dos mais populares para resolução de problemas lagrangeanos duais. Isto porque é extremamente simples e de fácil implementação, podendo ser descrito resumidamente da seguinte forma. 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.

Na relaxação #1, tem-se um conjunto de multiplicadores $\lambda \geq 0$, que 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 $\xi_1^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) = (\xi_{1_\ell}^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p))$ definidos abaixo:

$$\xi_{1_\ell}^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) = \sum_{t_p \in P} \delta_p \bar{f}_\ell^p - \Delta \bar{w}_\ell, \quad \forall \ell \in A. \quad (14)$$

Já na relaxação #2, tem-se um conjunto de multiplicadores (π, λ) , que 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 $\xi_2^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) = (\xi_{2_{ip}}^{\text{TGP}}(\bar{f}_\ell^p), \xi_{2_\ell}^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p))$ definidos abaixo:

$$\xi_{2_{ip}}^{\text{TGP}}(\bar{f}_\ell^p) = \sum_{\ell \in A^+(i)} \bar{f}_\ell^p - \sum_{\ell \in A^-(i)} \bar{f}_\ell^p - b_i^p, \quad \forall t_p \in P, \forall i \in N, \quad (15a)$$

$$\xi_{2_\ell}^{\text{TGP}}(\bar{w}_\ell, \bar{f}_\ell^p) = \sum_{t_p \in P} \delta_p \bar{f}_\ell^p - \Delta \bar{w}_\ell, \quad \forall \ell \in A. \quad (15b)$$

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

$$\lambda^{k+1} = \lambda^k + t^k d_1^k(\bar{w}_\ell, \bar{f}_\ell^p) \quad (16)$$

ou

$$(\pi, \lambda)^{k+1} = (\pi, \lambda)^k + t^k d_2^k(\bar{w}_\ell, \bar{f}_\ell^p). \quad (17)$$

em que t^k é um escalar positivo (tamanho do passo), enquanto $d_1^k(\bar{w}_\ell, \bar{f}_\ell^p)$ e $d_2^k(\bar{w}_\ell, \bar{f}_\ell^p)$ são direções de subida gerada a partir de ξ_1^{TGP} e ξ_2^{TGP} , respectivamente. Neste trabalho utilizou-se a seguinte adaptação da *regra de Crowder* para obtenção de uma direção de busca:

$$d^k = \frac{\xi^k + \rho d^{k-1}}{1 + \rho}, \quad \text{com } \rho \text{ fixo é igual a } 0,7. \quad (18)$$

A convergência do método de subgradientes está intimamente relacionada com o tamanho do passo t^k . Neste trabalho, optou-se por utilizar uma abordagem tradicional:

$$t^k = \sigma^k \frac{\bar{L} - L^k}{\|\xi^k\|^2} \quad (19)$$

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 ($L_1^{\text{TGP}}(\lambda)$ ou $L_2^{\text{TGP}}(\pi, \lambda)$). Finalmente, $\|\xi^k\|$ é uma norma qualquer, em geral, a norma euclidiana do respectivo subgradiente (ξ_1^{TGP} ou ξ_2^{TGP}) na k -ésima iteração. O valor de σ^k deve ser atualizado durante o processo de busca. Primeiramente, tomou-se $\sigma^1 = 1$, 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} . Além disso, um procedimento simplificado foi utilizado para se obter uma estimativa da solução do problema lagrangeano dual:

$$\bar{L} = \frac{\kappa \hat{L}_{\min} + L^k}{2} \quad (20)$$

em que κ é um escalar pouco maior que um (por exemplo, $\kappa = 1,05$) e \hat{L}_{\min} é o valor do melhor limite superior encontrado até o momento.

5. Experimentos Computacionais

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 [Bertsekas and Tseng 1994]; enquanto que para o cálculo de caminhos mínimos foi implementado o algoritmo de Dijkstra utilizando *D-Heaps*.

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 iteração. Como os testes preliminares apresentaram um melhor desempenho quando todas as três heurísticas para inicialização de \hat{A}_0 foram utilizadas, optou-se por armazenar todas as soluções viáveis obtidas heurísticamente de modo a se reduzir o tempo total. 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.

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 requisições foram gerados randomicamente (5 instâncias para cada quantidade de requisições). Todas as requisições possuíam um valor de demanda unitário e a capacidade máxima transporte de um λ foi fixada em 63. Já o conjunto II contém 50 instâncias do **TGP** para redes irregulares com 6, 10, 15 e 20 nós. A Figura 3 exibe as topologias irregulares usadas para geração das instâncias de teste. 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 cuja topologia é exibida na Figura 3(e). Durante os testes, utilizou-se os seguintes parâmetros: $Max\ Iter = 1000$, $\epsilon_1 = 10^{-3}$ e $\epsilon_2 = \epsilon_3 = 10^{-10}$. Além disso, para se minimizar tanto o número de *lightpaths* quanto o comprimento dos mesmos, os custos estabelecidos foram os seguintes, em que k_ℓ representa o número da camada à qual pertence o arco ℓ :

$$c_\ell = \begin{cases} 1000 + k_\ell, & \text{se } \ell \in A^e \\ 10 + k_\ell, & \text{se } \ell \in A^o \end{cases}, \forall \ell \in A. \quad (21)$$

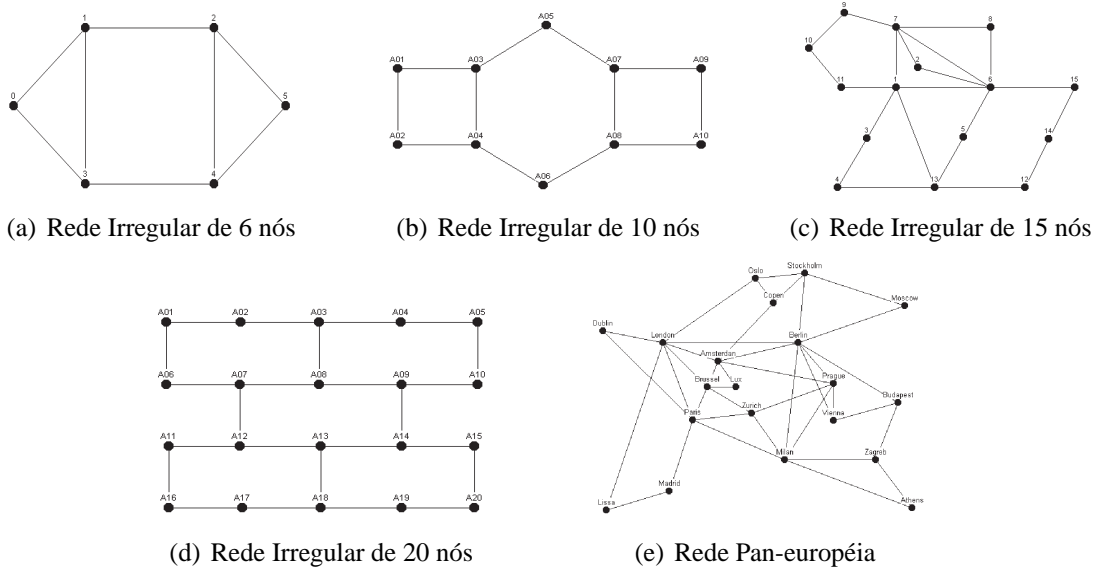
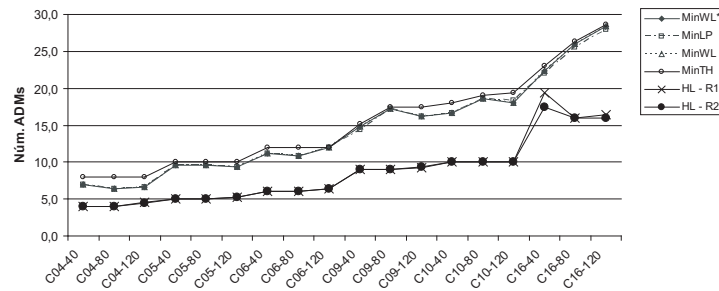


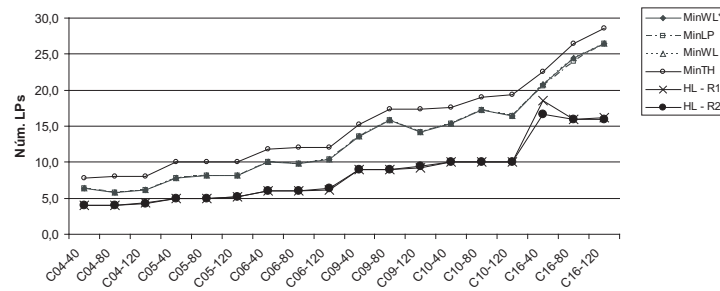
Figura 3. Redes com Topologia Irregular

Os resultados obtidos pela heurística (em um processador Pentium 4 de 1,8 GHz com 512 MB RAM) foram comparados com aqueles obtidos pela heurística proposta em [Zhu et al. 2003], sendo que os resultados obtidos pela heurística lagrangeana estão indicados na legenda pelas siglas **HL-R1** e **HL-R2** de acordo com a relaxação utilizada (relaxação #1 ou #2). Devido as diferenças na estrutura de custo utilizadas pelas políticas definidas em [Zhu et al. 2003] – **MinLP**, **MinWL** e **MinTH** – e as propostas apresentadas neste trabalho, a comparação entre elas deve ser realizada de forma indireta (através da política **MinWL*** – descrita em [Patrocínio Jr. 2005]). Em primeiro lugar, deve dizer que os resultados obtidos por **MinWL*** superam aqueles obtidos pelas demais políticas definidas [Zhu et al. 2003]. Tal constatação se baseia no fato de que o custo do projeto de rede virtual está diretamente relacionado ao número de *lightpaths* e de ADMs e que **MinWL*** obtém soluções com um número menor de *lightpaths* e de ADMs (comparada às políticas de [Zhu et al. 2003]). Por outro lado, se tais resultados forem comparados aos obtidos por **HL-R1** e **HL-R2**, verifica-se uma clara superioridade da abordagem lagrangeana em relação a heurística proposta em [Zhu et al. 2003].

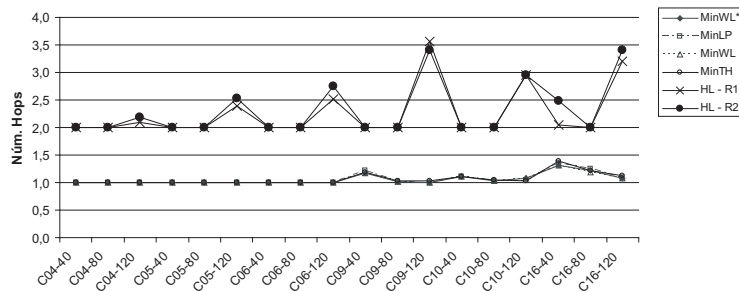
Os gráficos da Figura 4 apresentam uma comparação entre os valores médios obtidos pela heurística proposta em [Zhu et al. 2003] – de todas as políticas, inclusive **MinWL*** – e os resultados de **HL-R1** e **HL-R2** para as redes em anel. Observa-se que qualquer uma das abordagens lagrangeanas obtém valores 50% menores que os obtidos pela heurística proposta em [Zhu et al. 2003] tanto para o número de ADMs quanto para o número de *lightpaths* (LPs). Em verdade, as soluções obtidas pela heurística são 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 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 (ADMs e *lightpaths*). A heurística proposta em [Zhu et al. 2003], além de examinar uma única vez cada requisição, também se mostra incapaz de modificar as alocações já realizadas anteriormente; dificultando, assim, a



(a) Comparação do Número de ADMs



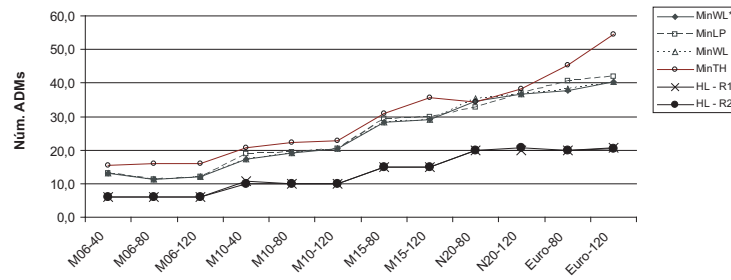
(b) Comparação do Número de LPs



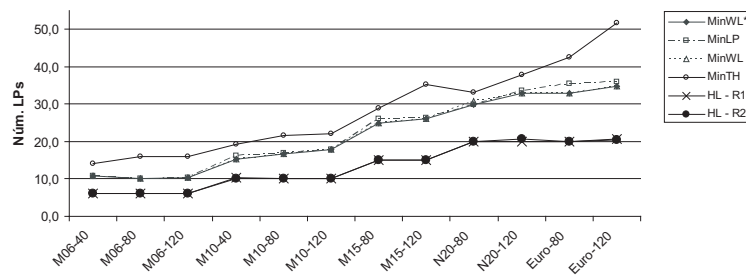
(c) Comparação do Número de Hops

Figura 4. Comparação dos Resultados Médios da Heurística Lagrangeana para Redes em Anel

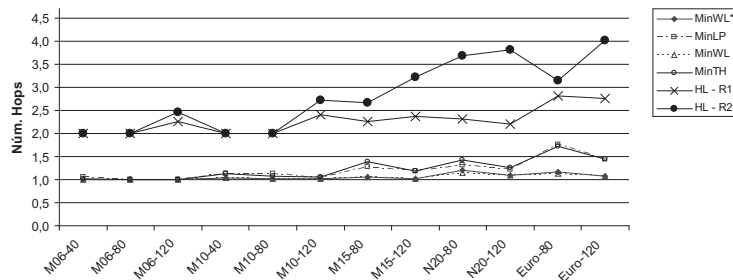
obtenção de melhores soluções. Além disso, os resultados obtidos pelas abordagens lagrangeanas podem ser melhor interpretados examinando-se o número médio de *hops* nas soluções obtidas por **HL-R1** e **HL-R2**. Em verdade, ao se examinar o gráfico da Figura 4(c), verifica-se que o número médio de “saltos” (*hops*) na topologia virtual obtida pelas abordagens lagrangeanas chega a ser de duas a três vezes maior que aquele obtido pela outra 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*) para rotear as requisições de tráfego (isto é, produzirem *lightpaths* maiores). Por outro lado, a heurística proposta em [Zhu et al. 2003] 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.



(a) Comparação do Número de ADMs



(b) Comparação do Número de LPs



(c) Comparação do Número de Hops

Figura 5. Comparação dos Resultados Médios da Heurística Lagrangeana para Redes Irregulares

Já os gráficos da Figura 5 apresentam uma comparação entre os valores médios obtidos pela heurística proposta em [Zhu et al. 2003] – inclusive com a política **MinWL*** – e os resultados de **HL-R1** e **HL-R2** para as redes irregulares. Novamente, os resultados obtidos por **HL-R1** e **HL-R2** para os valores médios de ADMs e de *lightpaths* superam (isto é, são menores) que os obtidos pela heurística proposta em [Zhu et al. 2003].

Mais ainda, analisando-se o gráfico da Figura 5(c), observa-se o mesmo padrão encontrado para as redes em anel. Contudo, para **HL-R2** o número médio de “saltos” (*hops*) chega a ser de quatro a cinco vezes maior que o obtido por **MinWL***, enquanto para que **HL-R1** esse número continua a ser de duas a três vezes maior. Este comportamento pode ser atribuído, pelo menos em parte, ao fato de que a relaxação #2 piora seu desempenho (isto é, a qualidade das soluções obtidas) quando aplicada a grandes redes irregulares. Isto se deve a forma como é resolvido o subproblema de roteamento de tráfego.

Na relaxação #2, esse subproblema é aproximado de maneira “simplista” através de um procedimento de resolução por inspeção, enquanto que na relaxação #1 a resolução do subproblema de roteamento é feita através do cálculo de vários caminhos mínimos. Essa estratégia, apesar de ser computacionalmente mais cara, apresenta melhores resultados em relação ao número de “saltos” (*hops*).

6. Conclusão

Neste trabalho 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 dois limites inferiores obtidos a partir de sua aplicação, procurou-se descrever os detalhes pertinentes à implementação de um procedimento de busca 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 de soluções.

Apresentou-se os resultados da aplicação da heurística lagrangeana para o **TGP** na resolução de mais de uma centena de instâncias do mesmo. Os resultados obtidos demonstram a superioridade das relaxações #1 e #2 em relação a outras abordagens. Além disso, a relaxação #1 parece ser mais adequada a resolução de grandes instâncias de redes irregulares. De modo a se avaliar qualitativamente as soluções obtidas pelas implementações da heurística lagrangeana – **HL-R1** e **HL-R2**, implementou-se e testou-se a heurística proposta em [Zhu et al. 2003]. Os resultados obtidos apontam para a superioridade da abordagem lagrangeana, em especial da relaxação #1, apesar de seu maior custo computacional.

Referências

- Bertsekas, D. and Tseng, P. (1994). RELAX-IV: a faster version of the RELAX code for solving minimum cost flow problems. Technical report, Massachusetts Institute of Technology, Cambridge, Massachusetts.
- Dutta, R. and Rouskas, G. (2000). A survey of virtual topology design algorithms for wavelength routed optical networks. *Optical Networks Magazine*, 1(1):73–89.
- Dutta, R. and Rouskas, G. (2002). On optimal traffic grooming in WDM rings. *IEEE Journal on Selected Areas in Communications*, 20(1):110–121.
- Patrocínio Jr., Z. (2005). *Planejamento de Topologia Virtual com Combinação de Tráfegos em de Baixa Velocidade Redes Óticas Multiplexadas por Divisão de Comprimento de Onda*. PhD thesis, Depart. Ciência da Computação - UFMG, Brazil.
- Wang, J., Vemuri, V., Cho, W., and Mukherjee, B. (2001). 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.
- Zhu, H., Zang, H., Zhu, K., and Mukherjee, B. (2003). A novel generic graph model for traffic grooming in heterogeneous WDM mesh network. *IEEE/ACM Transactions on Networking*, 11(2):285–299.
- Zhu, K. and Mukherjee, B. (2002). Traffic grooming in an optical WDM mesh network. *IEEE Journal on Selected Areas in Communications*, 20(1):122–133.