

MARCELO DE OLIVEIRA LIMA

***METODOLOGIA PARA O PROJETO COMPLETO
DE REDES ÓPTICAS COM TOPOLOGIA EM
HIERARQUIA***

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Engenharia Elétrica.

Orientador:

Prof. Dr. Marcelo Eduardo Vieira Segatto

Co-orientador:

Prof. Dr. Elias Silva de Oliveira

Vitória – ES

25 de setembro de 2010

Ficha catalográfica...

MARCELO DE OLIVEIRA LIMA

**METODOLOGIA PARA O PROJETO COMPLETO DE REDES
ÓPTICAS COM TOPOLOGIA EM HIERARQUIA**

Dissertação submetida ao programa de Pós-Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para a obtenção do Grau de Mestre em Engenharia Elétrica.

Aprovada em 26 de julho de 2010.

COMISSÃO EXAMINADORA

Dr. Marcelo Eduardo Vieira Segatto
Departamento de Engenharia Elétrica - UFES
Orientador

Dra. Rosane Bodart Soares
Departamento de Engenharia Elétrica - UFES
Examinador Interno

Dr. Renato Tannure Rotta Almeida
Instituto Federal de Educação C&T do Espírito Santo -
IFES
Examinador Externo

Dr. Carlos Renato Lisboa Francês
Faculdade de Engenharia de Computação - UFPA
Examinador Externo

“A mente que se abre a uma nova idéia jamais voltará ao seu tamanho original.”

Albert Einstein

Dedico esta dissertação a minha esposa Rafaela.

Agradecimentos

Tenho muito a agradecer a muitas pessoas e tenho certeza de que não conseguirei expressar totalmente, mas vale a tentativa.

Agradeço aos meus orientadores Segatto e Elias, com quem durante muitos anos esse trabalho foi construído, obrigado pela confiança e oportunidade.

Ao Renato Tannure pelas várias excelentes correções de textos nesse longo período.

Aos professores do PPGEE que em alguns momentos contribuíram para a abertura da mente. E aos colegas com quem partilhei bons momentos.

Aos meus gerentes durante essa longa trajetória pela flexibilidade de horário.

Aos meus familiares que me acolheram em Vitória e desde o tempo de graduação deram suporte e incentivo de grande valor.

Aos meus pais, Itamar e Tiana, a quem agradeço profundamente por terem me proporcionado bases sólidas. Começar bem faz uma grande diferença.

Aos meus irmãos, Filipe e Fabio. Difícil descrever, obrigado pelo companheirismo e amizade. Biu, sua dedicação absurda foi decisiva.

A minha esposa Rafaela, obrigado pelo apoio, incentivo e carinho. As incontáveis horas no computador deram resultado.

Agradeço aos meus amigos do MUR pelo incentivo, amizade e orações.

Agradeço finalmente a Deus, simplesmente pela direção, sustento e força.

Resumo

Este trabalho apresenta uma metodologia para o projeto físico e lógico de redes ópticas de comunicação com topologia de malhas hierárquicas. São determinadas as topologias lógica e física, além do roteamento e designação de comprimentos de onda, em função da localização geográfica dos nós da rede. A metodologia proposta consiste de três etapas que integram uma meta-heurística, inferência estatística e um modelo de programação linear inteira-mista. Na primeira um algoritmo genético define a estrutura hierárquica da rede óptica. Em seguida, um procedimento estatístico obtém estimativas para parâmetros de interesse que serão usados para definir critérios de qualidade para o projeto, limitando as variáveis do modelo de programação matemática resolvido na última etapa. São apresentados resultados de experimentos com o objetivo de validar a eficiência desta formulação quanto ao desempenho computacional e também com relação à qualidade das soluções, tendo como base de comparação limites inferiores para as métricas a serem otimizadas.

Palavras Chave: Redes ópticas, Otimização combinatória.

Abstract

In this work we present a methodology for logical and physical design of optical networks based on mesh hierarchical topologies. The proposed method can, simultaneously, design the logical and physical topologies, and perform routing and wavelength assignment. It is based on three steps: the metaheuristics, the statistical inference and the model of mixed-integer linear programming. The first one genetic algorithm defines the hierarchical structure of the optical network. Then, a statistical procedure to obtain estimate parameters of interest that will be used to define quality criteria for the project, limiting the model variables for the mathematical programming. We present results of experiments in order to validate the efficiency of this formulation on the computational performance and also regarding the quality of solutions, based on comparison of lower bounds for the metrics to be optimized.

Keywords: Optical networks, Combinatorial optimization.

Sumário

Lista de Figuras

Lista de Tabelas

1	Introdução	13
1.1	Principais Contribuições	16
1.2	Organização do Texto	17
2	Apresentação do Projeto de Redes Ópticas	18
3	Trabalhos Anteriores	23
3.1	Topologia Física	25
3.2	Topologia em Hierarquia	27
4	Modelo TWA	28
5	Metodologia para o Projeto Completo de Redes Ópticas em Hierarquia	37
5.1	Estrutura Hierárquica	41
5.2	Qualidade da Estrutura Hierárquica	43
5.3	Qualidade do Projeto das Sub-Redes	45
5.3.1	Congestionamento e Tráfego Retransmitido	46
5.3.2	Custo Físico	49
5.3.3	Restrições Adicionadas ao TWA	49
6	Resultados	51

6.1	Composição das Sub-Redes	51
6.2	Estatísticas de Congestionamento e Tráfego Retransmitido	55
6.3	Resolução das Sub-redes com TWA	61
7	Conclusões	67
	Publicações	70
	Referências Bibliográficas	72
	Apêndice A – Algoritmos Genéticos	75
	Apêndice B – Implementação de Algoritmos Genéticos	78

Lista de Figuras

2.1	Multiplexação e Demultiplexação em um sistema WDM.	19
2.2	WXC com duas portas de entrada e duas portas de saída.	19
2.3	Quatro sub-problemas se fundem em VTD e RWA.	22
4.1	Dois sub-problemas se fundem no TWA	29
4.2	Representação gráfica da notação associada aos nós da rede.	30
4.3	Representação gráfica de um componente topológico.	31
4.4	Exemplo da interpretação dos componentes topológicos.	32
5.1	Exemplo de topologia de rede em hierarquia.	37
5.2	Resumo da metodologia de projeto.	40
5.3	Ilustração do processo de decisão do genético.	44
6.1	Evolução do processo de busca do genético.	53
6.2	Histograma da distância média relativa às soluções visitadas pelo genético. . . .	54
6.3	Histograma do tráfego de acesso relativo às soluções visitadas pelo genético. . .	55
6.4	Convergência da média e desvio padrão em função do tamanho da amostra. . .	56
6.5	<i>Lower bounds</i> para o congestionamento.	59
6.6	<i>Lower bounds</i> para o tráfego retransmitido.	59
6.7	Resultados para tráfego retransmitido e de acesso ao <i>backbone</i>	65
A.1	Exemplo de uma solução para um TSP de quatro cidades.	76
A.2	Anel de 6 nós	77

Lista de Tabelas

6.1	Estrutura hierárquica para a rede de 30 nós. Em negrito, os nós do <i>backbone</i> . . .	53
6.2	Identificação das sub-redes.	57
6.3	Legenda para as tabelas 6.4, 6.5 e 6.6	57
6.4	Resultado das estatísticas para o congestionamento.	58
6.5	Resultado das estatísticas para o tráfego retransmitido.	58
6.6	<i>Upper bounds</i> calculados para congestionamento e tráfego retransmitido. . . .	61
6.7	Legenda para a tabela 6.8.	64
6.8	Resumo dos resultados com TWA.	64
A.1	Matriz de Adjacências de um anel	77

1 *Introdução*

O desenvolvimento de metodologias de projeto de redes ópticas de comunicações se intensificou após o advento da tecnologia de multiplexação por comprimento de onda (WDM - *Wavelength Division Multiplexing*). Além de possibilitar a implementação de vários canais de comunicação na mesma fibra óptica em diferentes comprimentos de onda, esta tecnologia permite a implementação de redes com roteamento de tráfego por comprimentos de onda (WRON - *Wavelength Routed Optical Networks*). As vantagens desse tipo de rede decorrem de sua infra-estrutura flexível, com elevada capacidade e confiabilidade na transmissão de dados.

No projeto de WRON técnicas de otimização são largamente empregadas e as soluções propostas fazem uso de métodos exatos e heurísticas. Na literatura, o projeto completo de WRON é dividido em dois sub-problemas (MUKHERJEE, 2006), que são denominados: VTD (*Virtual Topology Design*) e RWA (*Routing and Wavelength Assignment*). Esta divisão em sub-problemas é colocada como necessária para que se possa lidar com a extrema complexidade computacional das formulações de programação matemática que abordam o projeto completo de WRON. No entanto, em (LIMA et al., 2009; LIMA, 2010) foi apresentada uma modelagem para o projeto completo de WRON, denominada TWA (*Traffic over Wavelength Assignment*), capaz de tratar desde a escolha da topologia física da rede até a definição da topologia virtual, incluindo a distribuição de tráfego, a definição das rotas físicas e a designação e o roteamento de comprimentos de onda. Este modelo possui um reduzido número de variáveis e restrições, mesmo se comparado a modelos que resolvem apenas o RWA (JAUMARD; MEYER; THION-GANE, 2004).

Em uma WRON a conexão entre os nós da rede é realizada através da criação de caminhos ópticos. Nestas redes a topologia lógica ou virtual, se distingue da topologia física (RAMSWAMI; SIVARAJAN, 1996), ela é formada por um conjunto de rotas transparentes, onde o tráfego é transportado sem processamento e roteamento eletrônico. A topologia física é dita simétrica quando os enlaces físicos são compostos por pares de fibras utilizadas em sentidos opostos, configurando uma ligação bidirecional. Uma rede é chamada opaca se os caminhos ópticos coincidirem com os enlaces físicos da rede, desta forma exigindo a conversão do domínio

óptico para o eletrônico e posterior retransmissão. Em outro extremo estão as redes totalmente ópticas (WRAN - *Wavelength Routed All-Optical Networks*) onde existem caminhos ópticos interligando todos os nós, sem exigir qualquer passagem pelo domínio eletrônico, porém este não é um projeto eficiente em termos de recursos ópticos necessários aos nós da rede. As redes ópticas semi-transparentes, objeto de estudo neste trabalho, oferecem um bom compromisso entre a quantidade de recursos requeridos na camada óptica, principalmente em número de caminhos ópticos, e de recursos requeridos na camada eletrônica, basicamente a capacidade de processamento para comutação de tráfego (MUKHERJEE, 2006). Em tais redes, parte do tráfego é transportada de maneira totalmente transparente entre os pares de nós interligados diretamente por caminhos ópticos. Para os outros pares de nós, somente é possível o transporte de tráfego através de rotas formadas por mais de um caminho óptico em sequência. Neste caso, o tráfego deve ser processado nos nós intermediários de sua rota fonte-destino para que se efetue sua retransmissão pelo próximo caminho óptico.

As metodologias de projeto de WRON mais estudadas na literatura consideram topologias em malha para as redes ópticas. Existem também trabalhos que exploram redes com topologias hierárquicas, mais precisamente as implementadas na forma de anéis hierárquicos (SHI; FONSEKA, 1995; SEGATTO et al., 2006; SILVA et al., 2007). Contudo, até onde possuímos conhecimento, o projeto de topologias hierárquicas em malha ainda não foi explorado. A metodologia de projeto de WRON proposta considera que existe uma rede central na hierarquia superior que será chamada de *backbone*, cujos nós pertencem cada um a uma sub-rede de hierarquia inferior, chamada *cluster*. O objetivo principal da metodologia proposta neste trabalho é viabilizar o projeto completo de redes com topologias hierárquicas em malha. Considera-se como dados de projeto a disposição geográfica dos nós da rede e as demandas de tráfego esperadas entre os mesmos.

Uma simplificação adotada é que as ligações entre os *clusters* e o *backbone* serão opacas, ou seja, todos os componentes de uma ligação lógica estão em um mesmo nível da hierarquia. Deste modo, o *backbone* e os *clusters* podem ser vistos como redes independentes, exceto pelo tráfego que passa de um nível hierárquico para outro, chamado de tráfego de acesso. Este tráfego pode ser acumulado nos nós de acesso ao *backbone*, de modo que sua distribuição e roteamento podem ser feitos paralelamente em cada sub-rede.

Assim, para definir a estrutura hierárquica basta separar os nós da rede em conjuntos, que formarão cada *cluster*, e escolher em cada um qual será o nó de acesso ao *backbone*. Por sua vez, este último será formado pelo conjunto de nós de acesso, também chamados de super-nós (SEGATTO et al., 2006; SILVA et al., 2007). Portanto, com a possibilidade de agregar e atribuir o tráfego de acesso apenas aos super-nós, que passam a ser sua fonte e destino dentro das sub-

redes, temos que cada uma delas pode ser projetada separadamente. Destaca-se que o *backbone* será considerado como uma sub-rede, desta forma, o termo sub-rede se refere indistintamente ao *backbone* e aos *clusters*. Ao se considerar que os super-nós são opacos, diminui-se o alto custo computacional que seria esperado ao se formular o problema de projeto completo de uma WRON considerando todos os seus nós simultaneamente.

Outra estratégia adotada foi definir parâmetros de qualidade satisfatórios para as variáveis de interesse, de modo que qualquer solução viável encontrada seja suficiente para garantir a qualidade do projeto das sub-redes. As variáveis de interesse em questão são o congestionamento da rede, o total do tráfego retransmitido, o custo de instalação da rede física, o número de comprimentos de onda e o número total de transceptores utilizados (RAMASWAMI; SIVARAJAN; SASAKI, 2009; ALMEIDA et al., 2006; KRISHNASWAMY; SIVARAJAN, 2001; LIU; TOBAGI, 2008).

Ao se projetar redes WDM com roteamento de tráfego por comprimento de onda, devemos buscar uma solução que tanto minimize e distribua o tráfego alocado aos caminhos ópticos (MURTHY; GURUSAMY, 2002), como também minimize o atraso médio de pacotes na rede, entre outras opções, como por exemplo buscar a minimização da quantidade de tráfego processado eletronicamente nos nós (ALMEIDA et al., 2006). O congestionamento da rede é definido como a quantidade de tráfego alocado ao caminho óptico mais carregado da rede. Na fase de projeto, o congestionamento pode ser entendido como uma limitação de capacidade. A escolha dos enlaces de fibra óptica que deverão conectar os nós e compõem a topologia física da rede também necessita ser realizada com foco na otimização de recursos. Dessa forma, minimizar o custo físico da rede é mais um objetivo a ser considerado no projeto.

A metodologia proposta foi desenvolvida a partir do objetivo de resolver o projeto completo de redes ópticas para topologias de malhas hierárquicas. Os nós da rede são agrupados em sub-redes a partir de uma implementação de um Algoritmo Genético. A modelagem escolhida para o projeto da rede foi o TWA, que é aplicado separadamente para as sub-redes. Foram adotados critérios para garantir a qualidade da solução, sendo representada por parâmetros de projeto definidos. Utilizou-se indução estatística (NETO, 2002) para estimar a média e o desvio padrão para o congestionamento e tráfego retransmitido, a partir destas estatísticas foram definidos valores aceitáveis para esses parâmetros de projeto e comparados com *lower bounds* calculados. Para o custo físico da rede, definiu-se um indicador para mensurar o custo de instalação da rede física, que permite comparar o custo da solução encontrada com o de uma rede real conhecida.

A motivação para o estudo de topologia em hierarquia deve-se à facilidade relativa de implementação para um rede de tamanho elevado em comparação à uma topologia tradicional em

malha. Na prática esta é uma configuração encontrada em soluções para redes reais (THOMADSEN, 2005). Além disso, por questões estratégicas, existem situações onde é de interesse isolar os nós em sub-redes, devido por exemplo à expectativa de que falhas localizadas não afetem o desempenho geral da rede.

A sobrevivência (RAMASWAMI; SIVARAJAN; SASAKI, 2009) é um outro problema relacionado ao roteamento de tráfego em redes WDM e consiste em estabelecer rotas alternativas em caso de falhas no nós e também nos enlaces de fibra. Considera-se que a tolerância da rede a eventuais falhas está fora do escopo deste trabalho. Além disso, não será considerada qualquer capacidade de conversão de comprimentos de onda (ASSIS; WALDMAN; GIOZZA, 2005).

1.1 Principais Contribuições

Conforme foi mencionado, as topologias de WRON em malha vem sendo amplamente estudadas na literatura nas últimas décadas, as topologias na forma de anéis hierárquicos também foram foco de estudos (SHI; FONSEKA, 1995; SEGATTO et al., 2006; SILVA et al., 2007). Contudo, até onde possuímos conhecimento, o projeto de topologias em malhas hierárquicas ainda não foi explorado, sendo esta a principal contribuição deste trabalho.

Outra contribuição é a resolução completa do projeto de redes ópticas, tendo em vista que na literatura os sub-problemas que constituem este projeto são comumente tratados em separado, mesmo os trabalhos mais abrangentes (KRISHNASWAMY; SIVARAJAN, 2001; ASSIS; WALDMAN, 2004; TORNATORE; MAIER; PATTAVINA, 2007; LIMA, 2010) não cobrem todos os sub-problemas como será tratado aqui.

A própria metodologia proposta, como um todo, também representa uma contribuição. Ela permite a resolução das sub-redes em separado, graças a utilização do conceito de tráfego de acesso ao *backbone* e a decisão de tratar o roteamento do mesmo de forma opaca. Além disso, a formulação permite definir *a priori* critérios de qualidade satisfatórios para as principais métricas de interesse, permitindo aplicar o modelo matemático de otimização como uma ferramenta para escolha de uma solução viável, minimizando o elevado custo computacional tradicionalmente enfrentado nesta etapa.

A estratégia adotada para definir exigências de qualidade apresenta uma inovação, devido a aplicação de técnicas de indução estatística e a introdução de uma métrica para o custo da topologia física da rede, que permite comparação com redes reais conhecidas.

Com relação a utilização do modelo matemático TWA, neste trabalho supomos que a

topologia física é simétrica, o que representa uma adaptação do modelo que tira proveito desta hipótese, pois reduz o número de variáveis. Foram acrescentadas restrições ao modelo para a limitação do número de ligações lógicas e do tráfego retransmitido, além da nova forma de controle de custo físico proposta neste texto. Este trabalho traz a primeira aplicação envolvendo os custos de instalação da rede física e operação da rede, tal como foram propostos no modelo básico do TWA em (LIMA, 2010). Além disso, esta também é a primeira aplicação do TWA que utiliza sua capacidade de escolher a rede física conjuntamente com os demais aspectos do projeto.

1.2 Organização do Texto

O restante deste texto está organizado da seguinte forma: o capítulo 2 a seguir traz uma definição geral dos principais componentes do projeto de redes em estudo. No capítulo 3 são apresentados resumidamente os principais trabalhos da literatura nesta área de estudo. Na sequência, o capítulo 4 descreve o modelo TWA. A metodologia desenvolvida, foco deste trabalho, é apresentada no capítulo 5. A descrição dos experimentos computacionais realizados e os resultados obtidos estão no capítulo 6. Por fim são apresentadas as conclusões e perspectivas de trabalhos futuros.

2 *Apresentação do Projeto de Redes Ópticas*

Em uma rede óptica os dados são representados por sinais de luz. Mais especificamente, o princípio físico utilizado para comunicação são fótons tendo como meio de propagação fibras ópticas, diferentemente das redes tradicionais onde os dados são representados por sinais elétricos transmitidos por cabos. Algumas das vantagens são a menor atenuação dos sinais e a quase ausência de interferência entre sinais em uma mesma fibra, além da capacidade de transmissão em altíssima velocidade.

Partindo dos princípios de comunicação em fibras ópticas, a largura de banda disponível é enorme, com possibilidade de transmissão de dados em elevadas taxas, porém somente uma fração desta capacidade é efetivamente utilizada devido às limitações eletrônicas impostas pelos equipamentos que tem acesso as fibras. Uma tecnologia que melhora consideravelmente este panorama é a Multiplexação por Comprimento de Ondas (WDM - *Wavelength Division Multiplexing*), que envia simultaneamente um conjunto de sinais de luz em diferentes comprimentos de onda em uma mesma fibra, sendo que cada um desses comprimentos de onda corresponde a um usuário do enlace operando com uma taxa eletrônica de comunicação.

A figura 2.1 fornece uma visão conceitual da multiplexação e demultiplexação em um sistema WDM. Diferentes sinais de luz com estreitas larguras de banda, partindo de fontes distintas, são combinados para ocupar e aproveitar uma maior banda. No receptor os sinais são separados e direcionados para seus destinos.

Um componente básico das redes com roteamento por comprimento de onda, conhecido como *Wavelength Crossconnect* (WXC) ou *Wavelength Routing Switch*, tem a função de permitir a conexão (óptica) de qualquer canal de entrada (comprimento de onda) em uma porta de entrada (fibra óptica) com outra porta qualquer de saída ou a retirada do canal para processamento. A figura 2.2 ilustra conceitualmente um WXC.

Um mecanismo comum para transmissão em sistemas ópticos é constituído pelo uso de *lasers* semicondutores como fontes de luz. Os transmissores usados em redes WDM geralmente

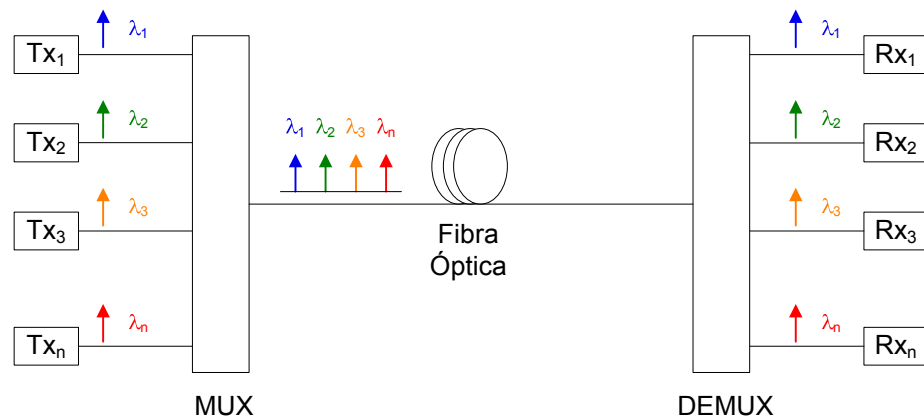


Figura 2.1: Multiplexação e Demultiplexação em um sistema WDM.

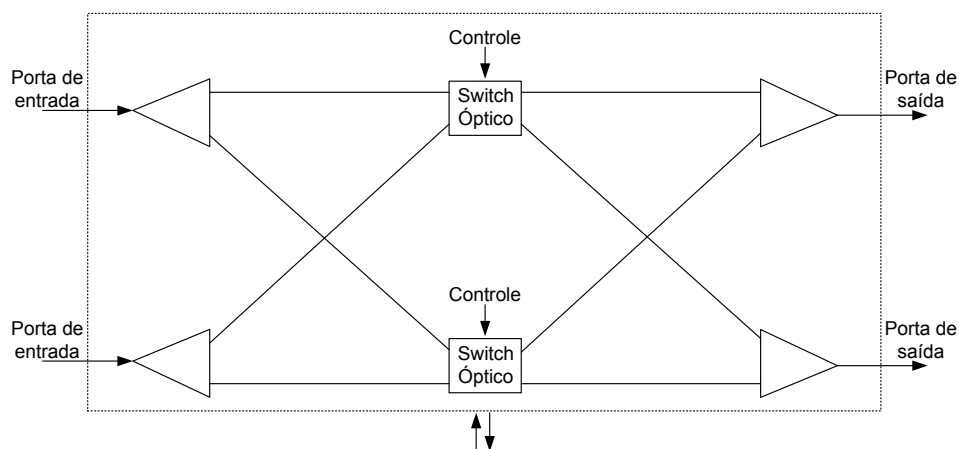


Figura 2.2: WXC com duas portas de entrada e duas portas de saída.

requerem a possibilidade de sintonia em diferentes comprimentos de onda.

Conforme descrito em (MURTHY; GURUSAMY, 2002), uma rede com roteamento por comprimento de onda (WRON - *Wavelength Routed Optical Networks*) consiste em WXCs, nós roteadores, interconectados por enlaces ópticos ponto-a-ponto formando uma topologia. Cada nó terminal é conectado a um WXC através de uma fibra óptica. A combinação de um nó terminal com seu correspondente WXC é referida, de forma geral, como nó (da rede). Cada nó é equipado com um conjunto de transmissores e receptores, todos com capacidade de sintonia de comprimento de onda.

Os dados são enviados de um nó para outro usando uma rota contínua, com um único comprimento de onda, chamada de caminho óptico, não requerendo qualquer conversão eletro-óptica ou armazenamento em fila nos nós intermediários. Este processo é conhecido como roteamento por comprimento de onda. Destaca-se que os nós intermediários roteiam os caminhos ópticos no domínio óptico usando seus WXC. Os nós terminais de um caminho óptico o acessam por meio de transmissores (início) e receptores (fim), de forma genérica chamados de tranceptores, que são sintonizados para o comprimento de onda no qual o caminho óptico

opera. Cada caminho óptico é associado a uma rota física e a um comprimento de onda. O requisito de que um mesmo comprimento de onda deve ser usado ao longo de todos os enlaces que compõem um caminho óptico é conhecido como restrição de continuidade de comprimento de onda. Dois caminhos ópticos não podem assumir o mesmo comprimento de onda em uma fibra, este requisito é conhecido como restrição de atribuição de comprimentos de onda distintos. No entanto, dois caminhos ópticos podem usar o mesmo comprimento de onda se eles percorrem conjuntos disjuntos de enlaces físicos, essa propriedade é conhecida como reutilização de comprimentos de onda.

Uma rede óptica com roteamento por comprimento de onda pode ser utilizada no roteamento de pacotes, para tanto, um conjunto de caminhos ópticos, representando a camada óptica da rede, são definidos sobre a sua camada física através da configuração dos WXC's em cada nó. Dessa forma, um pacote proveniente de um nó fonte, pode ser transmitido até um nó intermediário através de um caminho óptico, ser processado eletronicamente e posteriormente retransmitido para um caminho óptico subsequente, formando uma rota até alcançar seu nó de destino.

Se uma rede WDM entrega informação diretamente no domínio óptico para todos os pares de nós, ela é conhecida com uma Rede Totalmente Óptica (WRAN - *Wavelength Routed All-Optical Network*). Uma vantagem deste tipo de rede é evitar atrasos e perdas com conversões eletro-ópticas, processamento eletrônico e filas. A principal desvantagem está na limitação tecnológica e financeira em estabelecer caminhos ópticos entre todos os pares de nós da rede.

Os conjuntos de nós e enlaces físicos (fibras ópticas) constituem a topologia física. Enquanto o conjunto de caminhos ópticos definidos sobre a topologia física formam a topologia virtual ou lógica. Na topologia virtual, os nós correspondem aos próprios nós da rede e as arestas aos caminhos ópticos. O grau lógico de entrada de um nó na topologia virtual é definido como o número de caminhos ópticos que incidem no nó, enquanto o grau lógico de saída como o número de caminhos ópticos que partem do mesmo. Os graus lógicos de entrada e saída de um nó são limitados pelo número de receptores e transmissores ópticos disponíveis no nó, respectivamente. E ainda, considerando W o número de comprimentos de onda disponíveis por enlace físico e D_{in}^f e D_{out}^f os graus físicos de entrada e saída de um nó, respectivamente. Em uma topologia virtual o grau lógico de entrada (D_{in}^v) deste nó pode ser no máximo $W \times D_{in}^f$ e seu grau lógico de saída (D_{out}^v) pode ser no máximo $W \times D_{out}^f$.

Para o roteamento de tráfego pela topologia virtual, se dois nós estão conectados por uma aresta na topologia virtual, uma demanda pode ser transmitida de um nó a outro no domínio óptico, sem requerer qualquer conversão eletro-óptica em nós intermediários. Neste caso, a de-

manda é roteada em um salto virtual (lógico). Um WXC é usado para direcionar (no domínio óptico) a demanda que chega ao nó intermediário para um enlace de saída sem processamento de informação no nó. Uma demanda requer processamento eletrônico em um nó em duas situações: a demanda precisa ser destinada a um nó que não está conectado diretamente ao seu nó fonte através de um caminho óptico ou a demanda precisa ser recebida pelo seu nó destino. Uma demanda chegando a um nó no domínio óptico é convertida para sua forma eletrônica e armazenada. Caso o nó de destino desta demanda seja outro, ela é convertida novamente para o domínio óptico e transmitida por um enlace de saída. Por outro lado, se ela for destinada para o nó atual, ela é entregue eletronicamente para a camada superior da rede.

O projeto e planejamento de redes é realizado através de métodos distintos de acordo com o tipo de tráfego considerado, especificamente com relação a natureza estática ou dinâmica. No caso de tráfego estático, nosso foco de estudo, é assumido a priori uma determinada matriz de demanda de tráfego, representando a quantidade de tráfego que deve ser transferido entre os pares de nós da rede. Considera-se essas demandas como sendo fixas para fins de planejamento, podendo basear-se em levantamentos históricos ou mesmo estudos estimativos. Não serão consideradas aqui a possibilidade de bloqueio de pacotes ou outros tipos de perdas na transmissão. Portanto, é assumido que todo o tráfego da rede será devidamente enviado e recebido.

Uma rede óptica é transparente quando não existe regeneração eletrônica dos caminhos ópticos durante o seu percurso fim-a-fim, enquanto uma rede óptica é opaca quando cada caminho óptico é regenerado em todos os nós pelo qual transita na rede. Usando redes ópticas semitransparentes, é possível alcançar uma performance muito próxima aos das redes opacas em termos de bloqueio de novas requisições, porém com grande economia nos custos, e menos complexidade do que uma rede completamente óptica. Em suma, redes semitransparentes, foco do presente trabalho, oferecem o melhor dos domínios óptico e eletrônico sem comprometer as principais características de cada uma dessas tecnologias (STERN; ELLINAS; BALA, 2008).

O projeto de WRONs deve levar em conta seus custos de implementação e operação, que podem ser colocados, resumidamente, em função dos recursos de transmissão requeridos na camada óptica e a capacidade de processamento e armazenamento dos roteadores eletrônicos. Para tanto, técnicas de otimização são largamente empregadas e as soluções propostas fazem uso de métodos exatos e heurísticas, separadamente ou em conjunto. Na literatura (RAMASWAMI; SIVARAJAN; SASAKI, 2009), o projeto de WRON é dividido em quatro sub-problemas, que são denominados: roteamento de tráfego (TR - *Traffic Routing*), projeto da topologia lógica (LTD - *Logical Topology Design*), roteamento de comprimentos de onda (WR - *Wavelength Routing*) e alocação de comprimentos de onda (WA - *Wavelength Assignment*). Aos quatro subproblemas citados, podemos acrescentar a escolha da topologia física da rede, que neste

contexto significa definir os enlaces de fibra óptica que devem ser estabelecidos entre os nós da rede.

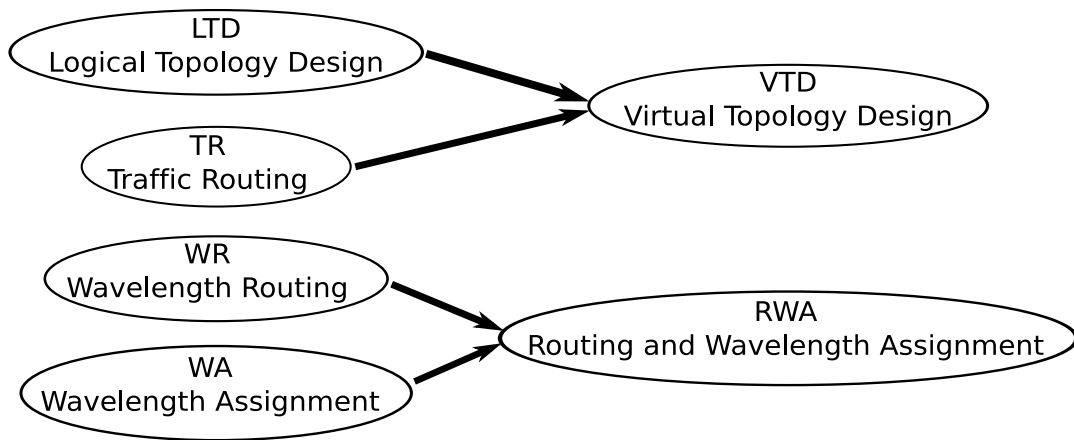


Figura 2.3: Quatro sub-problemas se fundem em VTD e RWA.

Tradicionalmente, os dois primeiros sub-problemas são associados, bem como os dois últimos, compondo, respectivamente, os conhecidos problemas de VTD (*Virtual Topology Design*) (RAMASWAMI; SIVARAJAN; SASAKI, 2009) e RWA (*Routing and Wavelength Assignment*) (ZANG; JUE; MUKHERJEE, 2000). Isto está ilustrado na figura 2.3.

Em resumo, o VTD pode ser descrito como a definição da topologia virtual da rede e posterior distribuição do tráfego pela mesma. Tendo como dados de entrada, a matriz de demandas de tráfego e o grau lógico considerado. Duas das principais métricas controladas ou otimizadas pelo VTD são o congestionamento e o tráfego retransmitido pelos nós (ALMEIDA et al., 2006). O congestionamento representa o total de tráfego alocado ao caminho óptico mais carregado da rede. O congestionamento pode ser entendido como a capacidade de transmissão dos caminhos ópticos, pois para se atender à demanda de tráfego, é necessário que essa capacidade seja no mínimo igual ao congestionamento calculado. O tráfego retransmitido pelos nós é uma característica das redes semitransparentes, pois é definido como a soma das demandas de tráfego que chegam e saem dos nós intermediários das rotas de comunicação entre os pares de nós da rede. Esse tráfego possui impacto direto sobre a capacidade de processamento eletrônico dos nós, e dessa forma sobre o seu custo, pois ele representa as demandas recebidas, processadas eletronicamente e encaminhadas novamente para o nó seguinte da rota.

O RWA (ZANG; JUE; MUKHERJEE, 2000) pode ser descrito como a definição das rotas físicas dos caminhos ópticos e posterior alocação de comprimentos de onda aos mesmos. Possui como dados de entrada uma matriz de requisições, que pode ser entendida como a própria topologia virtual, além da topologia física da rede. O objetivo principal é utilizar o menor número possível de comprimentos de onda para atender a todas as requisições.

3 *Trabalhos Anteriores*

Este capítulo faz um apanhado geral dos trabalhos encontrados na literatura que abordam o projeto de redes ópticas. O problema de projetar uma rede óptica pode ser formulado como um problema de programação inteira mista (MILP), sendo definida uma métrica de interesse a ser otimizada. Esse problema já foi amplamente estudado, muitas vezes recorrendo-se ao uso de métodos heurísticos para resolvê-lo, sendo conhecidamente NP-difícil. As diferentes abordagens partem de considerações específicas sobre as demandas de tráfego, tipo de topologia, capacidade de conversão de comprimentos de onda, métrica a ser otimizada, entre outras. Quando as demandas de tráfego são assumidas como sendo estáticas, caso adotado neste trabalho, todas as demandas a serem transmitidas entre os pares de nós da rede são conhecidas *a priori* e fixas. Neste caso, o objetivo normalmente é a minimização de algum recurso da rede, tendo como exemplos: número de comprimentos de onda utilizados, capacidade dos canais de comunicação, número de trancetores e processamento eletrônico. Com uma frequência menor, também são encontrados trabalhos que abordam o projeto da topologia física da rede, ou seja, a definição dos pares de nós que possuirão enlaces de fibra óptica interligando-os diretamente, neste caso de estudo o objetivo em geral é a minimização de um parâmetro de custo de instalação da rede, que normalmente varia entre o número de enlaces necessários ou o comprimento total de enlaces a serem instalados, existindo também casos onde se considera métricas como o número de comprimentos de onda ou de trancetores previstos.

Os trabalhos encontrados na literatura em geral resolvem parte dos subproblemas que constituem o projeto de uma WRON, dessa forma, as descrições apresentadas a seguir ressaltam as diferentes abordagens existentes.

O projeto de uma topologia virtual foi formulado como um problema de otimização em (MUKHERJEE et al., 1996). Os autores formularam o problema através de um modelo de otimização não linear. A função objetivo considerava a minimização do atraso na transmissão e do máximo fluxo em um enlace, sendo este último conhecido como o congestionamento da rede. Os autores subdividem o problema em quatro subproblemas: 1) determinação da topologia lógica; 2) roteamento dos caminhos ópticos sobre a topologia física; 3) alocação de comprimentos

de onda às rotas; 4) roteamento das componentes de tráfego (pacotes) na topologia lógica. Nos experimentos apresentados, os autores consideram apenas os subproblemas 1 e 4. Para a realização de experimentos, foi desenvolvido um método que emprega a meta-heurística *Simulated Annealing*. Posteriormente na literatura, o problema representado pelos subproblemas 1 e 4 conjuntamente passou a ser conhecido como *VTD (Virtual Topology Design)*, possuindo variações na função objetivo considerada. Em (RAMASWAMI; SIVARAJAN, 1996) os autores formularam uma modelagem MILP para resolução do *VTD* também com o objetivo de minimizar o congestionamento. Não existe restrição quanto ao número de comprimentos de onda utilizados. Nesta abordagem a topologia física torna-se irrelevante para o projeto da topologia lógica, pois não é resolvido o *RWA (Routing and Wavelength Assignment)*. O tráfego retransmitido na rede foi assumido como função objetivo em (ALMEIDA et al., 2006), enquanto o congestionamento era controlado através de uma restrição na formulação matemática utilizada. Conforme será detalhado no capítulo 5, a metodologia desenvolvida no presente trabalho limita o congestionamento e o tráfego retransmitido na rede baseando-se em exigências de qualidade estimadas por meio de inferência estatística.

Em (KRISHNASWAMY; SIVARAJAN, 2001) é proposta uma modelagem MILP que minimiza congestionamento em redes sem conversores de comprimentos de onda. Segundo os autores, esta formulação não é computacionalmente tratável, sendo métodos heurísticos propostos. O modelo MILP é relaxado e executado iterativamente por 25 vezes usando um plano de corte. As variáveis que representam a topologia virtual e os percursos físicos são arredondadas, enquanto uma heurística de alocação de comprimentos de onda é aplicada para atribuir comprimentos de onda individualmente aos caminhos ópticos. O tráfego é roteado pela topologia virtual utilizando uma formulação linear (LP) consistindo somente das restrições de tráfego do MILP relaxado. Uma desvantagem desse método, é que supondo que existam W comprimentos de onda disponíveis em cada fibra, o MILP relaxado obtém uma solução que satisfaz esta restrição. No entanto, na execução do algoritmo de alocação de comprimentos de onda, que é aplicado posteriormente e obtém soluções sub-ótimas, não há garantia de uma alocação de comprimentos de onda com sucesso, respeitando o limite W . Como resultado, o método não retorna necessariamente soluções viáveis para todos os casos. O modelo proposto é dos mais abrangentes da literatura, sendo que a única etapa do projeto de WRON não resolvida é a escolha da topologia física. Assim como no presente trabalho, a formulação permite a criação de múltiplos caminhos ópticos entre um mesmo par de nós da rede.

Um modelo MILP foi apresentado em (ASSIS; WALDMAN, 2004). Este trabalho propõe um algoritmo iterativo, que faz uso de programação linear, para resolver os problemas VTD e RWA de forma integrada. A topologia lógica é escolhida com a heurística *HLDA (RA-*

MASWAMI; SIVARAJAN, 1996), e esse resultado é fixado no modelo proposto. O modelo é resolvido para encontrar solução para as demais variáveis. Por se tratar de um modelo MILP de alto custo computacional, a resolução é interrompida depois de um tempo pré determinado. A função objetivo adotada foi o número total de saltos nas rotas físicas. A estratégia foi passar ao modelo limitações para as métricas importantes, de modo que as soluções viáveis encontradas fossem satisfatórias. Essa abordagem torna-se útil dada a abrangência do modelo proposto, onde métricas dos sub-problemas do projeto de uma WRON podem ser controladas. Todavia o alto custo computacional do modelo proposto inviabiliza sua aplicação para redes de grande porte. As redes testadas tinham 6 e 12 nós. No presente trabalho esta estratégia de definir exigências de qualidade para métricas de interesse e a heurística *HLDA* também serão empregadas, conforme está apresentado no capítulo 5.

3.1 Topologia Física

Em (BANNISTER; FRATTA, 1990) foi desenvolvida uma formulação para o projeto de redes WDM e analisados os princípios deste tipo de rede. O autor constatou a ocorrência de interferências entre o projeto das topologias virtual e física, considerando a qualidade da topologia virtual relacionada à parâmetros de performance da rede. A qualidade da topologia física foi relacionada ao custo de instalação, representado pelos comprimentos dos enlaces de fibra óptica. Adotou-se como objetivo minimizar o atraso médio de pacotes. Assim como em nosso trabalho, o autor utiliza uma matriz de demanda de tráfego que especifica a taxa, por exemplo em pacotes por segundo, na qual o tráfego é transmitido entre os pares de nós da rede, ao invés de requisições entre os pares de nós. O foco principal é a escolha das melhores topologias física e virtual possíveis mediante um conjunto de parâmetros de rede estabelecidos. Não é resolvida a alocação de comprimentos de onda.

O trabalho (TORNATORE; MAIER; PATTAVINA, 2007) trata de métodos exatos para o planejamento e otimização de redes WDM multi-fibras, ou seja, foi contemplada a criação de mais de um enlace de fibra entre um par de nós, diferentemente do presente trabalho. O autor propõe uma formulação para o problema chamada de *source formulation*, nela todo o fluxo é agregado em relação ao nó de origem. Esta formulação é equivalente à clássica, denominada *flow formulation*, porém permite uma redução relevante no número de variáveis e restrições, representando uma redução no tempo computacional durante a execução. O modelo *TWA* que será apresentado no próximo capítulo possui semelhanças conceituais quanto a agregação de tráfego. Com relação a conversão de comprimentos de onda, os casos extremos são tratados, quando todos os nós possuem capacidade de converter todos os comprimentos de

onda, e quando nenhum nó possui capacidade de conversão, sendo exigida a restrição de continuidade. O trabalho propõe a otimização da topologia virtual de uma rede física multi-fibra, com o objetivo de minimização de custo. O número de fibras por enlace necessárias para suportar uma matriz de tráfego pré-estabelecida é especificamente a variável a ser minimizada, tendo como dado de entrada o número de comprimentos de onda por fibra.

O projeto de redes totalmente ópticas com roteamento por comprimento de onda (WRAN) foi estudado em (LIU; TOBAGI, 2008). O autor o abordou a necessidade de decisão entre a instalação de mais fibras ou o aumento do número de comprimentos de onda utilizados. Ambas alternativas podem ser empregadas para suportar novas conexões, no entanto, uma eleva o custo de enlaces e a outra o dos nós, este último representado pelos equipamentos usados nos mesmos. Foi apresentado um algoritmo que busca soluções que minimizam o custo total para a topologia física, para tanto, empregou-se uma formulação matemática para o problema, além de um algoritmo heurístico. Segundo o autor, na instalação de uma WRAN é importante avaliar alternativas visando minimizar o capital investido, sendo o custo total dividido entre enlaces e nós. O custo dos enlaces está relacionado à instalação das fibras para interligar os nós da rede e é uma função do comprimento total dos enlaces. O custo relativo aos nós, é representado pelos roteadores de comprimentos de onda e é função do número de comprimentos de onda utilizados em cada enlace. Seria possível formar uma topologia física usando uma quantidade mínima de fibra a partir de uma árvore geradora mínima (*Minimum Spanning Tree*), no entanto esta configuração implicaria em um custo elevado para os nós, pois seria necessário uma grande quantidade de comprimentos de onda. Em outro extremo, seria possível conectar todos os pares de nós através de enlaces diretos. Isto implicaria em um custo mínimo para os nós, porém um alto custo para os enlaces. Foi apresentada uma formulação matemática e nos experimentos foi analisada a relação entre o custo dos nós e dos enlaces visando fornecer aos projetistas de redes uma ferramenta para o projeto de redes com boa qualidade técnica e financeira. Os conceitos abordados são aplicáveis a WRON, desde que sejam feitas as devidas considerações relativas às especificidades decorrentes da necessidade de retransmissão.

O projeto de topologia física para WRAN também foi estudado em (XIAO; LEUNG; HUNG, 2001). No problema abordado, é conhecido o número de requisições entre os pares de nós e as especificações de custo para estabelecer um enlace físico entre cada par de nós, tendo como objetivo determinar a topologia de menor custo para a rede. É apresentada uma formulação do problema, provado que se trata de um problema NP-completo e desenvolvido um algoritmo chamado *Two-Stage Cut Saturation*. Em um primeiro momento, é relaxada a restrição de continuidade de comprimentos de onda e aplicado o princípio de *Cut Saturation* para determinar uma solução inicial para a rede. Em seguida, no segundo estágio, é inserida a

restrição de continuidade de comprimentos de onda e resolvido o RWA para atender às requisições sobre a rede gerada inicialmente. Além da escolha da topologia física, este estudo possui semelhança com a metodologia apresentada no capítulo 5 devido a resolução do problema por etapas.

Os autores de (GUAN; CHAN, 1997) abordam o projeto de topologias física e lógica, analisando diferentes tipos de arquiteturas de redes. O custo das soluções são avaliados pelo número de enlaces de fibras utilizados e por uma modelagem de custo relacionado aos OXCs. Foi demonstrada a relevância dos saltos lógicos com relação ao dimensionamento dos recursos de roteamento, além de analisada a influência do grau dos nós sobre o desempenho das soluções.

3.2 Topologia em Hierarquia

Foi proposta em (SHI; FONSEKA, 1995) uma arquitetura de hierarquia em anéis chamada *Hierarchical Self-Healing Rings* (HSHR) e o projeto dessas redes explorado. A arquitetura HSHR consiste de anéis em níveis diferentes, sendo que um anel de nível superior é usado para conectar os anéis imediatamente inferiores, de forma que cada anel em um determinado nível somente se comunica diretamente a um único anel de nível imediatamente superior. Um modelo geral de custo incorporando materiais e instalação foi usado. Mostrou-se que o método de enumeração, que encontra a configuração ótima para a rede, somente pode ser utilizado para redes de pequeno porte devido a complexidade. São apresentadas heurísticas para encontrar soluções de qualidade para a configuração dos HSHR. O roteamento e dimensionamento dos anéis também são considerados.

Conforme descrito em (PENG; CHEN; WU, 2008) uma modelagem de rede em hierarquia pode reduzir a complexidade do roteamento e alocação de comprimentos de onda (RWA) em redes WDM quando comparados a modelos para redes em malha. No entanto, a utilização de recursos pode ser ineficiente se a arquitetura em hierarquia não for adequadamente planejada. Partindo do grau de cada nó e do impacto da localização, neste trabalho foi proposto um método heurístico para construção de topologias em hierarquia para redes WDM visando a maximização da utilização dos recursos de rede. A topologia da rede em hierarquia é obtida pelo agrupamento de nós com seus respectivos nós LRA (*Local Routing Area Node*), sendo que a interconexão dos nós LRA constitui uma hierarquia de nível superior. Com uma solução focada no desempenho do RWA, foram observados os resultados para a quantidade de comprimentos de onda utilizados, assumindo que cada enlace de fibra possui quatro comprimentos de onda disponíveis. A rede utilizada nos testes possui 28 nós. Foram avaliados diferentes métodos de roteamento e analisada a taxa de bloqueio na rede.

4 *Modelo TWA*

O modelo TWA (*Traffic over Wavelength Assignment*) foi introduzido em (LIMA, 2010), onde também são apresentados diversos casos de uso, sua descrição está reproduzida parcialmente neste capítulo. É apresentada uma adaptação do modelo que foi utilizada nos experimentos computacionais do capítulo 6. O TWA é capaz de tratar desde a escolha da topologia física de uma WRON até a definição da topologia virtual, incluindo a distribuição de tráfego, a definição das rotas físicas e a alocação de comprimentos de onda. Ainda assim, possui um reduzido número de variáveis e restrições, se comparado a modelos que resolvem apenas o RWA, como os que são tratados em (JAUMARD; MEYER; THIONGANE, 2004). Na literatura o projeto completo, incluindo topologias física e lógica, foi modelado em (XIN; ROUSKAS; PERROS, 2003), possuindo uma complexidade elevada, que torna o uso de heurísticas uma exigência.

O TWA guarda semelhanças com alguns modelos conhecidos (RAMASWAMI; SIVARANJAN; SASAKI, 2009; TORNATORE; MAIER; PATTAVINA, 2007). Fazendo uso das definições dos sub-problemas que compõem o projeto de redes ópticas, conforme apresentados no Capítulo 2, nas modelagens para o WA (ZANG; JUE; MUKHERJEE, 2000), é designado um comprimento de onda a cada caminho óptico, considerando o seu percurso físico determinado pelo WR. A configuração dos caminhos ópticos, em termos de quantidade, fonte e destino, é obtida pela solução do VTD. Esta abordagem de modelos separados para VTD e RWA exige variáveis diferentes para as ligações lógicas, para suas rotas físicas e para a alocação de comprimentos de onda (ASSIS; WALDMAN, 2004).

O TWA propõe uma visão diferente. Tendo sido alocados comprimentos de onda entre pares ordenados de nós com variáveis específicas, estarão determinadas implicitamente, pelas restrições do modelo, as rotas físicas e as ligações lógicas entre esses pares de nós. Escrevendo todas as restrições do RWA e do VTD apenas em termos dessas variáveis de alocação de comprimentos de onda, não serão necessárias variáveis adicionais para determinar as rotas físicas e as ligações lógicas, o que simplifica o modelo e o torna computacionalmente mais tratável.

As restrições do TWA, em função das variáveis de alocação de comprimento de onda, de-

terminam a configuração e o roteamento dos caminhos ópticos, o que define as topologias física e lógica, além da alocação de comprimentos de onda. Restando apenas resolver a distribuição do tráfego. Esta última é tradicionalmente feita em função das variáveis de topologia lógica e roteamento de tráfego do VTD (RAMASWAMI; SIVARAJAN; SASAKI, 2009), retornando a matriz de topologia lógica e as requisições de tráfego designadas a cada um dos caminhos ópticos. A partir daí, a solução do VTD é fornecida como entrada para o RWA (ZANG; JUE; MUKHERJEE, 2000), na forma de uma matriz de requisições de tráfego associadas aos caminhos ópticos. Diferentemente do procedimento tradicional, o TWA possui restrições para a distribuição do tráfego que também são escritas em função das variáveis de alocação de comprimentos de onda. Na prática, isso elimina as restrições de distribuição de requisições de tráfego do RWA (ZANG; JUE; MUKHERJEE, 2000). Isto está ilustrado na figura 4.1.

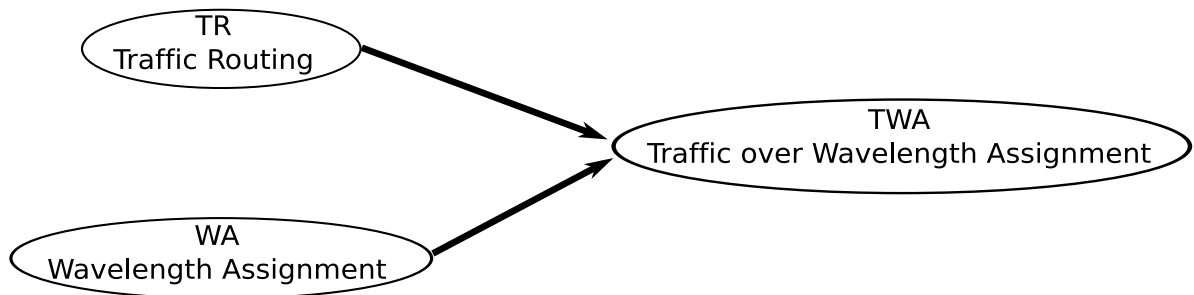


Figura 4.1: Dois sub-problemas se fundem no TWA

Assim sendo, as variáveis e restrições do TWA consistem em um modelo completo para o projeto de redes ópticas, pois considera todos os seus subproblemas de maneira integrada. Algumas vantagens foram incorporadas, a principal delas é que a distribuição do tráfego e seu roteamento são feitos com variáveis agregadas, de forma similar a modelos menos abrangentes encontrados na literatura (TORNATORE; MAIER; PATTAVINA, 2007; RAMASWAMI; SIVARAJAN; SASAKI, 2009). Outra característica é que ele naturalmente admite múltiplos caminhos ópticos e múltiplas fibras ópticas entre cada par de nós da rede, sem a necessidade de diferenciar cada ligação por uma variável de decisão diferente, como na abordagem utilizada anteriormente em (RAMASWAMI; SIVARAJAN; SASAKI, 2009).

A função objetivo adotada é a minimização dos custos de instalação e operação da rede, valendo-se da capacidade do modelo escolher também a topologia física da rede. Além disso, foi considerada a restrição de conservação dos comprimentos de onda ao longo do caminho óptico (ZANG; JUE; MUKHERJEE, 2000), ou seja, não se admite a conversão de comprimentos de onda na camada óptica da rede.

Notação 1. Para uma rede de N nós, os pares ordenados (m, n) , (s, d) e (i, j) indicam respectivamente ligações físicas, demandas de tráfego e ligações lógicas, com $m \neq n$, $s \neq d$ e $i \neq j$,

onde $m, n, s, d, i, j \in \{1, \dots, N\}$. O índice $w \in \{1, \dots, W\}$ representa comprimentos de onda, onde W é a quantidade limite de comprimentos de onda que podem ser usados. O índice $v \in \{1, \dots, N\}$ representa os nós da rede. Por sua vez, $[m, n]$ representa uma ligação física bidirecional.

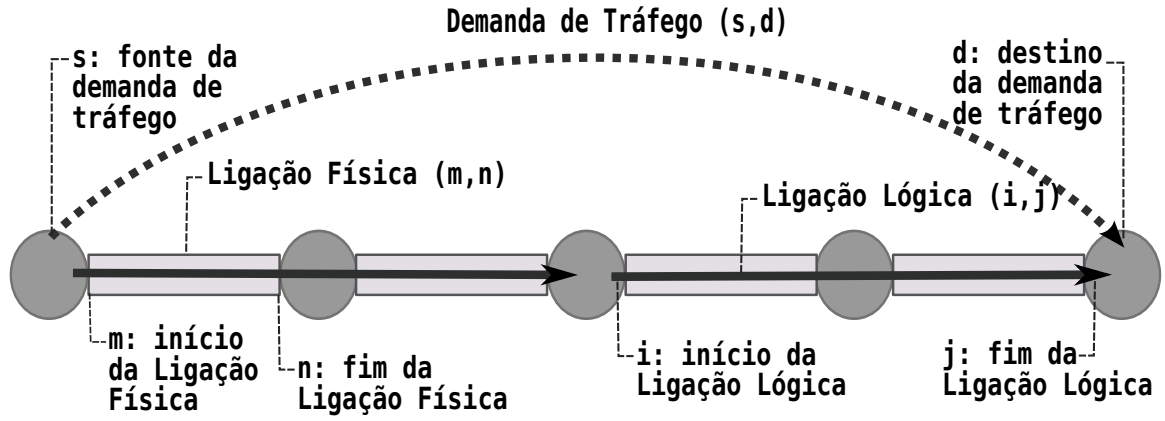


Figura 4.2: Representação gráfica da notação associada aos nós da rede.

A Figura 4.2 ilustra os diferentes escopos dos índices associados aos nós da rede, com relação às ligações físicas (m, n) , ligações lógicas (i, j) e demandas de tráfego (s, d) . Esta notação segue a convenção comumente utilizada em trabalhos anteriores (MUKHERJEE, 2006; RAMASWAMI; SIVARAJAN; SASAKI, 2009). É importante dizer que esta modelagem suporta múltiplas ligações físicas e lógicas entre cada par de nós, portanto, os pares (m, n) e (i, j) representam conjuntos de possíveis ligações físicas e lógicas, respectivamente. Esses conjuntos não serão explicitamente controlados, sendo esse um dos motivos da simplicidade do modelo.

Dados 1. Uma instância para o modelo TWA é definida por:

1. N = Número de nós da rede.
2. W = Máximo de comprimentos de onda em uma ligação física.
3. K = Máxima multiplicidade de ligações físicas entre cada par (m, n) .
4. Cap = Capacidade de tráfego de cada ligação lógica.
5. C_{mn} = Custo de uma ligação física bidirecional $[m, n]$.
6. T = Custo por unidade de fluxo.
7. P_{sd} = Demanda de tráfego, com origem s e destino d .
8. $A_s = \sum_d P_{sd}$ = Tráfego agregado pela origem s .
9. $Q_{sd} = P_{sd}/A_s$ = Fração de A_s correspondente à Demanda de tráfego P_{sd} .

A variável central do modelo, a partir da qual todas as demais serão definidas, chamada de componente topológico, é representada graficamente na Figura 4.3 e formalmente definida na Variável 4.0.1. Ela sozinha representa as topologias lógica e física, as rotas físicas das ligações lógicas e os comprimentos de onda utilizados.

Variável 4.0.1. *Seja $B_{iw}^{mn} = k \in \{0, \dots, K\}$, com $i \neq n$, um componente do conjunto das ligações lógicas com origem i e comprimento de onda w , que utilizam k ligações físicas entre os nós m e n .*

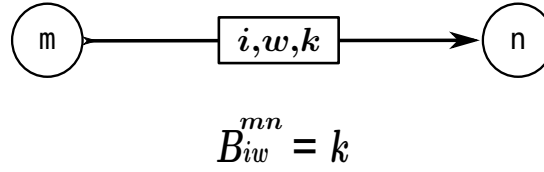


Figura 4.3: Representação gráfica de um componente topológico.

Considerando que $B_{iw}^{mn} = k$ para algum $k \in \{0, \dots, K\}$, existem k ligações lógicas originadas em i no comprimento de onda w , passando por k ligações físicas distintas entre o par de nós (m, n) . Conforme a terminologia utilizada neste trabalho daqui por diante, *um componente topológico $B_{iw}^{mn} = k$ é iniciado em m , incidente em n , com origem i , comprimento de onda w e valor k .*

Se $k > 1$, então há multiplicidade de ligações físicas entre o par de nós (m, n) , pois haveria interferência se houvessem duas ligações lógicas se propagando na mesma ligação física, com o mesmo comprimento de onda. Note que K limita apenas a multiplicidade das ligações físicas, pois se $K = 1$, B_{iw}^{mn} se torna uma variável binária, mas ainda podem haver múltiplas ligações lógicas entre um par (i, j) , utilizando rotas físicas distintas, ou ainda, comprimentos de onda diferentes em uma mesma rota física. Se $B_{iw}^{mn} = 0, \forall (i, w)$, então nenhuma ligação física entre o par de nós (m, n) é utilizada.

Na Figura 4.4, temos um exemplo de interpretação dos componentes topológicos, todos com origem no nó v_1 e com o mesmo comprimento de onda w . No item *d*) desta figura, o valor 2 do componente que liga os nós (v_1, v_2) é interpretado como duas ligações físicas entre esses nós, representadas no item *a*). No item *b*), vemos uma ligação lógica dupla entre os nós (v_1, v_3) , onde uma delas passa de forma transparente pelo nó v_2 , como indicado no item *c*). Note ainda que, no item *d*), há dois caminhos lógicos incidentes em v_2 mas apenas um iniciando. Isso indica que uma ligação lógica termina em v_2 , enquanto a outra segue adiante.

A indexação atribuída às variáveis B_{iw}^{mn} especificam apenas o nó i onde se iniciam as ligações lógicas representadas, sem deixar claro aonde elas terminam. Isto significa que estas variáveis

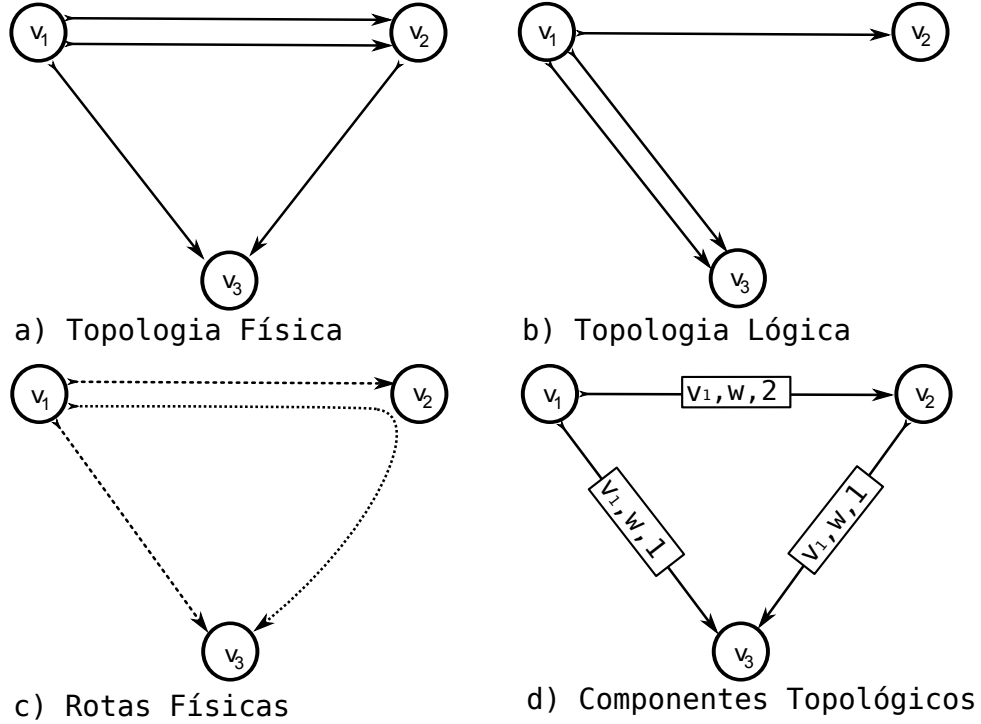


Figura 4.4: Exemplo da interpretação dos componentes topológicos.

agregam todas as ligações lógicas originadas em i com comprimento de onda w , que utilizam as ligações físicas entre o par (m, n) , independente do nó j em que terminam estas ligações lógicas. Esta técnica consiste em uma abordagem bastante conhecida para a representação de variáveis em problemas de distribuição de fluxo em redes. Em (TORNATORE; MAIER; PATTAVINA, 2007), este conceito de agregação de tráfego é aplicado como meio de simplificação do modelo, reduzindo substancialmente o número de variáveis dos problemas resultantes. No TWA, esta agregação cumpre o mesmo papel de simplificação, cabendo às restrições do modelo garantir implicitamente a terminação correta destas ligações lógicas agregadas nas variáveis B_{iw}^{mn} .

As Variáveis 4.0.2 e 4.0.3 completam as definições necessárias para apresentarmos a forma básica do modelo TWA.

Variável 4.0.2. *Seja $q_{sw}^{ij} \in [0, 1]$ a fração do fluxo originado em s , passando pelas ligações lógicas entre o par (i, j) com comprimento de onda w , onde $s \neq j$.*

Para resolver o sub-problema de roteamento de tráfego, é definida a Variável 4.0.2, que modela a fração de fluxo agregado para as demandas de tráfego. Elas são semelhantes às variáveis de fluxo agregado utilizadas em (RAMASWAMI; SIVARAJAN, 1996), todavia há duas diferenças. Uma delas é que aqui essas variáveis são normalizadas em função do tráfego agregado na origem (A_s), e são portanto uma fração deste. Essa modificação não é requerida pela modelagem, tendo apenas a função de facilitar a compreensão das restrições do modelo, que

ficam menos dependentes dos dados de entrada. A outra diferença é que o fluxo é separado por comprimento de onda, como se fossem W redes sem multiplexação sobrepostas. Isso facilita a interpretação das restrições do modelo, e também ajuda a mantê-lo mais simples.

Variável 4.0.3. *Seja $D_{mn} \in \{0, \dots, K\}$, com $m > n$, o número de ligações físicas bidirecionais entre o par de nós m e n .*

Apesar da topologia física ser determinada pelos componentes topológicos, para fins de controle do custo de instalação da rede física, são necessárias novas incógnitas. Para este fim, é definida a Variável 4.0.3, que registrará em D_{mn} a multiplicidade física determinada pelos componentes topológicos. Se $D_{mn} = 0$, não há ligações físicas entre o par (m, n) , mas se $D_{mn} = k$, para algum $k \in \{0, \dots, K\}$, existem k ligações físicas bidirecionais $[m, n]$.

Duas métricas importantes no projeto da redes ópticas são os custos de instalação e operação (MUKHERJEE, 2006). O custo de instalação C_{mn} é o custo associado a uma ligação física entre o par de nós (m, n) . O custo total de instalação é dado na equação 4.0.1. O custo de operação, é definido como o custo por unidade de fluxo e calculado na equação 4.0.2, influencia também no dimensionamento dos nós da rede.

$$CI = \sum_{mn} C_{mn} \cdot D_{mn} \quad (4.0.1)$$

$$TO = \sum_{sijw} T \cdot q_{sw}^{ij} \cdot A_s \quad (4.0.2)$$

O custo de operação pode ser dividido em duas partes: uma constante, formada pelas demandas de tráfego (equação 4.0.3), que necessariamente deverão ser roteadas; e outra variável, composta pelo tráfego adicional que é gerado, ou seja, o tráfego retransmitido (equação 4.0.4). A parte constante do custo de operação não influenciaria na função objetivo, por isso não será incluída em seu cálculo, dado na equação 4.0.5.

$$TOC = \sum_{sd} T \cdot P_{sd} \quad (4.0.3)$$

$$TOV = \sum_{sijw} T \cdot q_{sw}^{ij} \cdot A_s, \quad i \neq s \quad (4.0.4)$$

$$FO = CI + TOV \quad (4.0.5)$$

Outro ponto positivo dessas métricas é que minimizar o custo por unidade de fluxo é equivalente a minimizar o tráfego retransmitido na rede, o que por sua vez, equivale a minimizar o processamento eletrônico de tráfego dos nós da rede (ALMEIDA et al., 2006). Além disso, será necessária nesta modelagem uma restrição de limitação da capacidade das ligações lógicas (*Cap*), que equivale à limitar o congestionamento na rede. Assim, limitando a capacidade e minimizando o custo de operação, temos uma abordagem eficiente, quanto ao custo computacional, para controlar também o congestionamento e o processamento, importantes métricas no projeto da topologia lógica (ALMEIDA et al., 2006; RAMASWAMI; SIVARAJAN; SASAKI, 2009).

Se não for necessário ponderar o custo por unidade de fluxo, basta fazer $T = 1$, e se não for necessário considerar o custo total de instalação (*CI*), basta fazer $C_{mn} = 0$ para todo (m, n) . Deste modo seria simplesmente um modelo de minimização do processamento, com limitação do congestionamento (ALMEIDA et al., 2006).

A seguir é apresentada a formulação para o TWA considerada neste trabalho.

Função Objetivo

- Custo de Instalação e Operação

$$\text{Minimize: } \sum_{mn} C_{mn} \cdot D_{mn} + \sum_{sijw} T \cdot q_{sw}^{ij} \cdot A_s, \quad , \text{ com } i \neq s \text{ e } m > n \quad (4.0.6)$$

Restrições

- Continuidade de Comprimentos de Onda e Capacidade:

$$\sum_s q_{sw}^{iv} \cdot A_s \leq Cap \cdot \left(\sum_m B_{iw}^{mv} - \sum_n B_{iw}^{vn} \right), \quad \forall (i, v, w), \text{ com } i \neq v \quad (4.0.7)$$

- Topologia Física:

$$\sum_i B_{iw}^{mn} \leq D_{mn} \text{ se } m > n, \text{ ou } D_{nm} \text{ se } m < n, \quad , \quad \forall (m, n, w) \quad (4.0.8)$$

- Conservação de Fluxo:

$$\sum_{jw} q_{vw}^{vj} = 1, \quad \forall v \quad (4.0.9)$$

$$\sum_{iw} q_{sw}^{iv} - \sum_{jw} q_{sw}^{vj} = Q_{sv}, \quad \forall (s, v), \text{ com } s \neq v \quad (4.0.10)$$

Note que os componentes topológicos estão definidas em função de ligações físicas orientadas. Todavia, convencionamos que as ligações físicas estabelecidas serão bidirecionais.

Acumulando múltiplas funções, a Restrição 4.0.7 atua como restrição de continuidade de comprimentos de onda e limitação de capacidade. Ela garante a continuidade das rotas físicas, onde os componentes topológicos devem formar um caminho sobre a topologia física, conservando o mesmo comprimento de onda. Esses percursos não são controlados explicitamente; eles são garantidos pela conservação dos componentes topológicos nos nós intermediários, semelhante a uma restrição de conservação de fluxo (RAMAMURTHY et al., 1999).

Nela, a conservação dos percursos lógicos é feita da seguinte forma: a soma dos componentes das ligações lógicas iniciadas em um nó i no comprimento de onda w , partindo de um nó intermediário v , deve ser menor ou igual à quantidade recebida. A Restrição 4.0.7 é um conjunto de equações, onde cada uma trata de um par (i, j) em um comprimento de onda w . Portanto, a capacidade combinada das múltiplas ligações lógicas associadas ao par (i, j) é a capacidade de cada uma (Cap) multiplicada pelo número de ligações lógicas entre (i, j) no comprimento de onda w .

Com a finalidade de controlar pela função objetivo 4.0.6 a quantidade de ligações físicas definidas pelos componentes topológicos, a Restrição 4.0.8 acumula nas variáveis D_{mn} a multiplicidade determinada pelos componentes. Dado um par (m, n) , as equações dessa restrição são ainda separadas por comprimento de onda. Pois se todos os componentes topológicos alocados em (m, n) usarem o mesmo w , apenas uma ligação física será necessária. Se usarem comprimentos de onda diferentes, D_{mn} precisará atender ao maior desses componentes topológicos. Portanto, a restrição 4.0.8, minimiza a soma dos componentes topológicos em cada par (m, n) , por força do fator CI na função objetivo (equação 4.0.1).

A conservação de fluxo é assegurada pelas Restrições 4.0.9 e 4.0.10, que também garantem o envio e a entrega das demandas de tráfego. Elas são semelhantes às encontradas na modelagem agregada para o VTD (RAMASWAMI; SIVARAJAN; SASAKI, 2009). Cada par (i, j) é visto nas restrições de controle de fluxo como um único caminho, unindo todos os comprimentos de onda. Se o par representar na verdade múltiplas ligações lógicas, a diferença é que ele terá uma capacidade maior de receber tráfego, que é controlada pela Restrição 4.0.7. Deste modo, essas restrições funcionam da mesma forma que em (RAMASWAMI; SIVARAJAN, 1996).

A Restrição 4.0.9 garante que todo o tráfego originado em cada nó v seja emitido para a

rede, exigindo que a soma das frações de tráfego, em todos os comprimentos de onda, que iniciam na origem ($i = s = v$) seja igual a 1, ou seja, 100% do tráfego originado em v .

Por sua vez, a Restrição 4.0.10 garante que o tráfego emitido seja encaminhado através da rede e entregue no destino. Fixada uma origem de tráfego s , para cada nó intermediário v ($v \neq s$) a porção de tráfego que deve ser entregue é Q_{sv} . Ela é igual à soma do tráfego chegando por todos os comprimentos de onda w , vindo de qualquer nó intermediário i , subtraída da soma do tráfego partindo com destino a qualquer nó j , em qualquer comprimento de onda w . O tráfego não entregue em v continua seguindo seu caminho pela rede até seu destino, e deste modo é feita rastreabilidade do tráfego até sua origem.

O tráfego pode ser subdividido e transportado simultaneamente por mais de uma ligação lógica entre o par (i, j) , no comprimento de onda w . Neste caso, como as rotas terão o mesmo comprimento de onda, eles não compartilham ligações físicas ao longo do percurso. Mas essas rotas podem ainda não ser disjuntas, pois é possível compartilharem nós intermediários.

5 Metodologia para o Projeto Completo de Redes Ópticas em Hierarquia

O tipo de topologia para redes ópticas com roteamento por comprimento de onda mais estudado na literatura é a topologia em malha. No entanto, existem trabalhos que exploram redes com arquitetura em hierarquia, mais precisamente anéis em hierarquia (SHI; FONSEKA, 1995), (SEGATTO et al., 2006), (SILVA et al., 2007). Contudo, até onde possuímos conhecimento, a junção destes dois tipos de arquiteturas ainda não foi explorado no contexto de WRONs. Neste trabalho, decidiu-se por estudar uma arquitetura de malhas em hierarquia, onde existe uma sub-rede na hierarquia superior que será chamada de *backbone*, cujos seus nós pertencem cada um a uma outra sub-rede de hierarquia inferior. Cada sub-rede inferior é chamada de *cluster*. Estas sub-redes podem assumir arquiteturas em malha ou anel, de acordo com o grau lógico em estudo. Quando se trabalha com grau lógico 1, trata-se de um anel, caso contrário de uma topologia em malha. A figura 5.1 ilustra uma topologia hierárquica genérica.

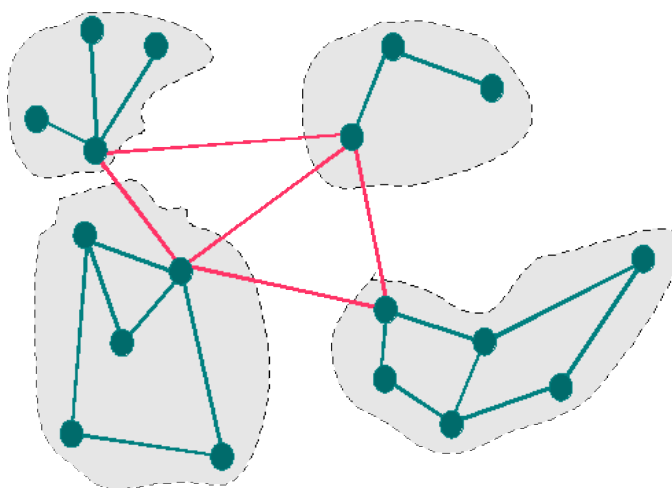


Figura 5.1: Exemplo de topologia de rede em hierarquia.

O objetivo principal deste trabalho é criar uma metodologia que viabilize o projeto completo de malhas hierárquicas, aproveitando o potencial da modelagem TWA, juntamente com

métodos heurísticos. Tomando como requisitos apenas um conjunto de nós de rede, com sua disposição geográfica e requisições de tráfego esperadas. O projeto completo abrange aqui, além das variáveis de decisão suportadas pelo modelo TWA, a escolha da estrutura hierárquica da rede.

Idealmente, o projeto completo de uma malha hierárquica poderia ser feito incluindo restrições no TWA que modelassem a hierarquia, semelhante ao que foi feito em (THOMADSEN, 2005). Todavia, com o objetivo de tratar do projeto de redes de grande porte, optamos por uma abordagem simplificada em termos de modelagem matemática, deixando a escolha da composição das sub-redes a cargo de uma heurística desenvolvida a partir de algoritmos genéticos. Conforme mencionado, as hierarquias projetadas neste trabalho terão apenas dois níveis, com um *backbone* central e *clusters* periféricos, de forma semelhante aos exemplos encontrados na literatura para anéis hierárquicos.

Uma simplificação adotada é que as ligações entre os *clusters* e o *backbone* serão opacas, ou seja, todos os componentes de uma ligação lógica estão em um mesmo nível da hierarquia. Deste modo, o *backbone* e os *clusters* podem ser vistos como redes independentes, exceto pelo tráfego que cruza de um nível hierárquico para outro, chamado de tráfego de acesso. Este tráfego pode ser acumulado nos nós de acesso ao *backbone* (como será mostrado na Seção 5.2), de modo que sua distribuição e roteamento podem ser feitos paralelamente em cada sub-rede.

Assim, para definir a estrutura hierárquica basta separar os nós da rede em conjuntos, que formarão cada *cluster*, e escolher em cada um qual será o nó de acesso ao *backbone*. Por sua vez, este último será formado pelo conjunto de nós de acesso, também chamados de super-nós (SEGATTO et al., 2006), (SILVA et al., 2007).

Dessas considerações, temos que cada sub-rede pode ser projetada separadamente, eliminando a necessidade de uma modelagem para toda a hierarquia.

A escolha da estrutura hierárquica, como será definida, é um Problema de Cobertura de Conjuntos (PCC), com uma função objetivo especial. A formulação clássica do PCC é um problema de programação linear inteira NP-completo (CORMEN et al., 2002). Portanto, abordagens heurísticas são adequadas para sua resolução. Neste trabalho utilizamos um algoritmo genético nesta etapa.

Tendo sido determinada a estrutura hierárquica, cada sub-rede é modelada com o TWA. Nesta fase, a estratégia adotada foi definir parâmetros de qualidade satisfatórios para as variáveis de interesse, de modo que qualquer solução viável encontrada já seria suficiente para garantir a qualidade do projeto.

As variáveis de interesse em questão são o congestionamento da rede, o total do tráfego retransmitido, o custo de instalação da rede física, o número de comprimentos de onda utilizados e o número total de transceptores utilizados. As duas primeiras são comumente tratadas pela literatura no projeto de topologias lógicas (VTD), sem considerar o projeto da rede física ou o roteamento e alocação de comprimentos de onda (RWA) (RAMASWAMI; SIVARAJAN; SASAKI, 2009).

Para estabelecer parâmetros de qualidade para congestionamento e tráfego retransmitido, são estimados estatisticamente as médias e desvios padrões de suas populações de soluções viáveis para o VTD, em cada sub-rede. Nessas inferências é garantida a representatividade das amostras obtidas, como será mostrado na seção 5.3.1.

Para o número de comprimentos de onda, foi utilizada uma estratégia adotada anteriormente em (LIMA et al., 2009). Tenta-se obter soluções com um comprimento de onda apenas e incrementa-se esse número enquanto não for possível encontrar soluções viáveis. Essa técnica deriva de uma característica do TWA, pois o número de variáveis no modelo é múltiplo do número de comprimentos de onda adotado.

Por sua vez, o custo de instalação da rede física é controlado garantindo que seja menor, proporcionalmente, que o custo de uma rede real bem conhecida, a NSFNET (RAMASWAMI; SIVARAJAN; SASAKI, 2009). A forma como é estabelecida essa proporcionalidade será descrita na seção 5.3.2. Considera-se como custo de instalação apenas a distância entre os nós da rede, que representa o comprimento dos enlaces de fibra instaladas entre os mesmos. O dimensionamento dos nós poderá ser feito com base no projeto obtido desta metodologia, que deverá demandar o mínimo possível de recursos da rede.

No projeto de topologias lógicas, o número de transceptores da rede normalmente define as instâncias do problema (RAMASWAMI; SIVARAJAN; SASAKI, 2009), pois está relacionado com o dimensionamento dos nós da rede. Essa abordagem também foi utilizada aqui, todavia, na literatura é comum definir o número exato de transceptores que cada instância possui, pelo produto entre grau lógico e o número de nós da rede, com níveis diferentes de limitação de tráfego para cada uma (KRISHNASWAMY; SIVARAJAN, 2001). Neste trabalho, apenas foi limitado o número de transceptores, permitindo que o modelo encontre soluções mais econômicas para este parâmetro.

As próximas seções são dedicadas a explicação detalhada dos principais aspectos que compõem as etapas da metodologia proposta. Uma visão geral do procedimento está apresentada na figura 5.2, que ilustra a existência de três etapas sequenciais.

Na primeira etapa, são fornecidos como dados as matrizes de demandas de tráfego e dis-

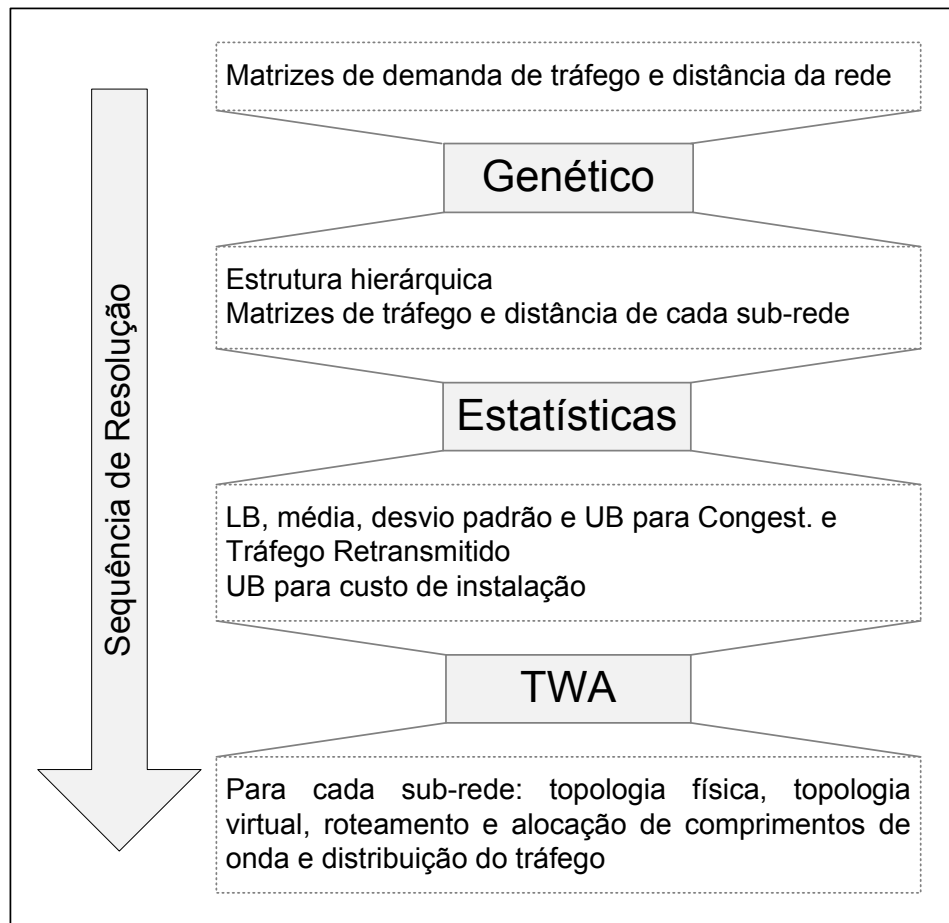


Figura 5.2: Resumo da metodologia de projeto.

tâncias, como dados de entrada do projeto. Um algoritmo genético faz a separação dos nós em sub-redes. Sua função objetivo minimiza o tráfego transferido entre as sub-redes e a distância entre os nós, ou seja, procura-se concentrar o tráfego no interior das sub-redes, e formá-las com nós geograficamente próximos. Um resultado secundário dessa fase são as matrizes de distância e tráfego para cada sub-rede.

Na fase intermediária, são obtidos os limites superiores (*upper bounds* - UB) para o congestionamento e o tráfego retransmitido na rede, por meio de inferência estatística baseada na média e desvio padrão dos resultados das avaliações de amostras de soluções viáveis para o VTD. Estes limites superiores são valores máximos admissíveis para o congestionamento e o tráfego retransmitido, impostos às variáveis do modelo de programação inteira-mista para garantir, com relação a estas métricas, a qualidade da solução a ser obtida na próxima etapa. Além dos UB, para validar os resultados, são também calculados limites inferiores teóricos (*lower bounds* - LB) para essas métricas. As formulações para os LB foram apresentadas por (LIMA, 2010), e os UB são calculados a partir da média e do desvio padrão. Por fim, também são calculados UB para o custo de instalação física de cada sub-rede.

A última etapa consiste na resolução do modelo TWA para cada sub-rede, respeitando os *bounds* (UB e LB) calculados e visando utilizar o menor número possível de comprimentos de onda. O resultado desta última fase fornece a solução do projeto, representado pelas topologias física e lógica, bem como a forma de roteamento e alocação dos comprimentos de onda, em função da distribuição das demandas de tráfego. Esses resultados são referentes às sub-redes em separado, porém, quando são agrupados, representam o resultado global para uma WRON com topologia de malhas hierárquicas.

5.1 Estrutura Hierárquica

A quantidade e o tamanho dos *clusters* é livre, sendo definidos apenas os tamanhos mínimo M_i e máximo M_a . A partir daí, para uma rede de n nós, ficam implicitamente definidas as quantidades máxima ($Q_a = \lfloor n/M_i \rfloor$) e mínima ($Q_i = \lfloor n/M_a \rfloor$) permitidas de *clusters*. Assim, uma rede de n nós e k *clusters*, é representada por dois vetores. Um vetor de tamanho k contendo os tamanhos dos *clusters* ($T = [t_1, \dots, t_k]$) e outro de tamanho n contendo uma permutação dos nós da rede ($P = [p_1, \dots, p_n]$). Onde o primeiro *cluster* é representado por $[p_1, \dots, p_{t_1}]$, o segundo por $[p_{(1+t_1)}, \dots, p_{(1+t_2)}]$, e o i -ésimo *cluster* é representado por:

$$[p_{(1+s_{[i-1]}), \dots, p_{(s_i)}], \text{ onde } s_i = t_1 + t_2 + \dots + t_i$$

Em cada *cluster*, o primeiro nó da sequência é interpretado como o nó de acesso ao *backbone*. Portanto, o *backbone* é representado por:

$$[p_1, p_{(1+t_1)}, \dots, p_{(1+s_{[k-1]})}]$$

Para a escolha da estrutura hierárquica da rede foi implementado um algoritmo genético com a biblioteca *Galib* (Apêndices A e B). Cabe ressaltar que para esta etapa poderia ter sido adotada outra meta-heurística ou procedimento de otimização combinatória, no entanto optou-se pelo algoritmo genético devido a experiência prévia do grupo de pesquisa que o autor faz parte, a relativa facilidade para implementação e desempenho satisfatório.

Para representar a hierarquia, é definido um cromossomo utilizando um vetor de tamanho $n + Q_a$, contendo uma permutação dos números inteiros positivos até esse valor: n números para representar os nós da rede mais Q_a números, candidatos a tamanho de *cluster*. A interpretação do cromossomo como uma hierarquia é feita da seguinte forma:

1. Separa-se o cromossomo em dois vetores, mantendo-se a ordem dos números. O primeiro, de tamanho n , contém os números menores ou iguais à n . Este é o vetor P , uma permutação dos nós da rede. O segundo vetor ($C = [c_1, \dots, c_{Q_a}]$) contém os números maiores que n e dará origem ao vetor T de tamanhos.
2. A partir de c_1 , iterativamente, faz-se $t_i = M_i + [c_i \bmod \Delta]$, enquanto $n - s_i \geq M_i$. Sendo $\Delta = (M_a - M_i)$ e s_i como definido acima. Ou seja, enquanto o número de nós restante não for menor que o mínimo admitido para formar outro *cluster*. Quando isso ocorrer, descarta-se o valor calculado para t_i e forma-se o último *cluster* com os nós restantes.

Note que o cálculo de t_i , como foi definido acima, leva o valor de c_i para um intervalo de possíveis tamanhos de *clusters*. Os valores que darão origem ao vetor de tamanhos T , foram definidos como números maiores que n para simplificar a codificação do cromossomo. Sendo este uma simples permutação de inteiros entre 1 e $n + Q_a$, sem uma estrutura particular, permitiu utilizar operadores genéticos pré disponíveis na biblioteca *Galib* e bem conhecidos na literatura (MICHALEWICZ; ZBIGNIEW, 1994). São eles os operadores genéticos do código *Ex26* da *Galib*, com a função de cruzamento: *Edge Recombination Crossover*.

Deve-se garantir que a soma de tamanhos possíveis ultrapasse n , o que pode não ocorrer dependendo dos valores de M_a , M_i e Q_a . Isso é garantido se a inequação a seguir for verdadeira:

$$n \leq \sum_{i=1}^{Q_a} M_i + [(n + i) \bmod \Delta]$$

Pelo fato dos operadores genéticos utilizados suportarem indivíduos formados por permutação de números inteiros, não são geradas soluções inviáveis. Isso contribui para a eficiência do algoritmo (GOLDBERG, 2002), pois evita visitar soluções que não contribuem para a melhora na evolução. Além disso, economizou-se o esforço em seu desenvolvimento, implementação e validação.

O algoritmo genético adotado foi o *GADemeGA*, também disponível na *Galib*, que mantém populações paralelas com probabilidade de migração a cada geração da evolução. Os parâmetros de configuração das funções da *Galib* foram mantidos em seus valores padrão e demonstraram desempenho satisfatório, mas para implementações mais robustas é recomendado que sua calibragem seja melhor estudada.

5.2 Qualidade da Estrutura Hierárquica

Além da codificação do cromossomo e da escolha das funções da *Galib* a serem utilizadas, o algoritmo genético se completa com a definição da sua função objetivo, que atribuirá uma métrica de qualidade para os cromossomos.

Dois parâmetros importantes no projeto de uma rede são os custos de instalação e operação da mesma (BANNISTER; FRATTA, 1990), (LIU; TOBAGI, 2008), (GUAN; CHAN, 1997). Na separação dos conjuntos de nós que formarão as sub-redes um fator que influencia fortemente no custo de instalação é a distância entre os nós. Se as sub-redes forem agrupados de modo que nós próximos façam parte do mesmo conjunto, haverá uma tendência de menor custo de instalação dos enlaces. Por esse motivo, a distância média entre os nós é uma boa métrica para minimizar o custo de instalação físico.

Não obstante, se dois nós que possuem grande quantidade de demanda de tráfego entre eles estiverem no mesmo *cluster*, eles deixarão de rotear pelo *backbone* esse forte fluxo. Se essa tendência se mantiver por toda a rede, com as maiores demandas representando tráfego doméstico dos *clusters*, é razoável supor que esta configuração poupará custo de operação da rede, pois estará reduzindo o tráfego de acesso ao *backbone* e também o retransmitido entre nós intermediários. Definimos tráfego de acesso como sendo todo aquele que cruza as hierarquias da rede, em ambos os sentidos, ou seja, se origina em uma sub-rede com destino a outra. Destaca-se que para efeito de resultados do projeto, ele é computado apenas nos super-nós que fazem a sua interface entre hierarquias distintas, pois nos demais nós ele é contabilizado indistintamente como tráfego retransmitido.

A figura 5.3 ilustra o processo de decisão do genético, onde ele busca agrupar os nós fisicamente mais próximos e ao mesmo tempo seleciona em cada grupo, ou sub-rede, um nó que será utilizado como interface entre as demais sub-redes, de forma a minimizar o tráfego de acesso. Esses nós de acesso estão destacados no interior de cada sub-rede.

Deste modo, a função objetivo do algoritmo genético foi definida como a soma da distância média entre os nós da hierarquia (M) com o total de tráfego de acesso ao *backbone* (T). Esses valores podem ser calculados sem que seja necessário conhecer a estrutura completa da rede, basta sabermos a estrutura hierárquica, como foi definida na seção anterior.

Como resultado do algoritmo genético, além da estrutura hierárquica da rede, também são derivadas as matrizes de demanda e distância de cada sub-rede. As matrizes de distâncias são apenas um subconjunto da original. Enquanto que as matrizes de demanda acumulam nos nós de acesso todo tráfego que deverá entrar no *cluster* e sair dele.

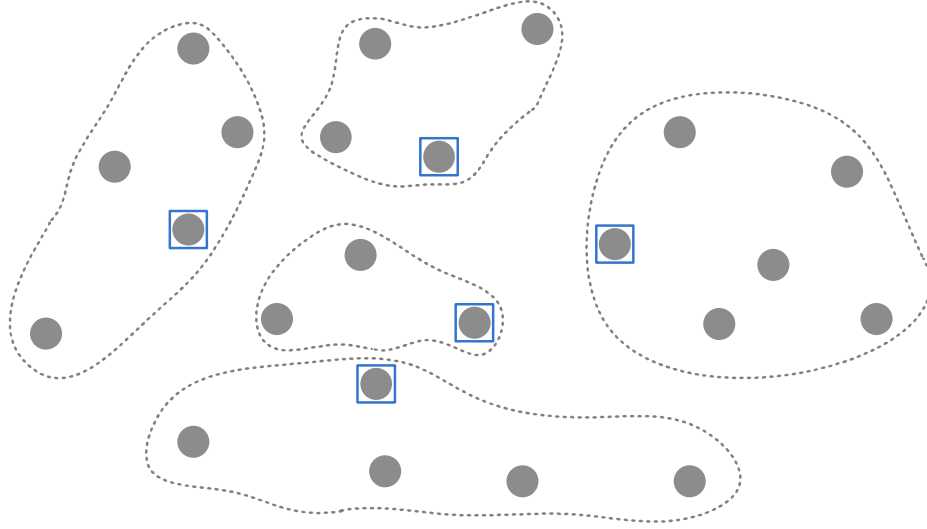


Figura 5.3: Ilustração do processo de decisão do genético.

A equação 5.2.1 define a fórmula de cálculo para M . Primeiro calcula-se a soma das distâncias entre os nós de cada grupo, *backbone* ($TDist_{bk}$) e *clusters* ($TDist_{cl}$), e a média desses valores é tomada como a distância média da hierarquia. Dependendo do estudo de caso em questão, pode ser de interesse ponderar esses termos, possivelmente atribuindo maior peso às distâncias do *backbone*. Na equação 5.2.1 esses pesos ($PesoDist_{bk}$ e $PesoDist_{cl}$) estão apresentados multiplicando seus respectivos termos de distância. O denominador no cálculo de M é a soma da quantidade de termos contidos nas matrizes de distâncias do *backbone* ($NumDist_{bk}$) e dos *clusters* ($NumDist_{cl}$).

$$M = \frac{TDist_{bk} \cdot PesoDist_{bk} + TDist_{cl} \cdot PesoDist_{cl}}{NumDist_{bk} + NumDist_{cl}} \quad (5.2.1)$$

Para calcular o tráfego de acesso em um super-nó, soma-se todo o tráfego originado dentro do seu *cluster* cujo destino são nós de outros *clusters*, mais o tráfego da condição inversa. Ou seja, que origina-se em outros *clusters* e possui destino no interior do *cluster* a que pertence o referido super-nó. Acumulando-se essa soma para todos os super-nós, temos o total de tráfego de acesso da rede T .

A equação 5.2.2 mostra que o total de tráfego de acesso também pode ser ponderado ($PesoTrafAcesso$) de acordo com o interesse do estudo realizado.

$$T = TrafAcesso \cdot PesoTrafAcesso \quad (5.2.2)$$

Nos experimentos que serão apresentados no próximo capítulo, as entradas das matrizes

de demanda de tráfego e distância possuem a mesma ordem de grandeza. Podendo representar respectivamente a taxa de transferência de pacotes em *Gbits/s* e centenas de quilômetros de distância por exemplo. Fica claro que neste caso M , por ser uma média, tem ordem de grandeza bem inferior à de T , que é uma soma simples. Portanto, para garantir o equilíbrio entre as duas métricas, um teste piloto deve ser feito para estabelecer uma proporção que deixe M e T com a mesma ordem de grandeza. Feito isso, também é possível atribuir pesos para cada um, conforme for o interesse do estudo de caso em questão. A equação 5.2.3 ilustra a fórmula de cálculo da função objetivo do genético (*fitness*), ilustrando o uso de um fator de calibragem (C_a) a fim de tornar M e T compatíveis em termos de ordem de grandeza.

$$Fitness = M \cdot C_a + T \quad (5.2.3)$$

O critério de parada adotado para o algoritmo genético foi a convergência da função objetivo, baseada na estagnação da mesma. Quando deixar de haver melhora na evolução por um dado número de gerações, o algoritmo genético é interrompido. A eficiência deste critério foi analisada nos experimentos computacionais.

5.3 Qualidade do Projeto das Sub-Redes

Para cada sub-rede, são definidos conjuntos de parâmetros de qualidade para o seu projeto. Esses parâmetros devem ser considerados satisfatórios para cada instância, representando uma qualidade elevada para a rede, pois na última etapa do procedimento eles serão fornecidos para o modelo TWA a fim de se obter uma solução viável que atenda a todos esses parâmetros.

Distinguindo as instâncias pelo grau lógico, para o congestionamento nas sub-redes e o total de tráfego retransmitido nas mesmas, são analisadas estatisticamente as populações de soluções para o VTD. Pois essas são métricas comumente tratadas nesse problema (RAMASWAMI; SIVARAJAN; SASAKI, 2009). Dessa análise, serão estabelecidas as limitações para o congestionamento e para o tráfego retransmitido. Além disso, também será definida uma limitação para o custo de instalação da rede física, que é independente da instância.

Na seção 5.3.1 será explicado o processo de estimação das estatísticas para congestionamento e tráfego retransmitido. O critério de limitação do custo de instalação está descrito na seção 5.3.2. Por fim, para que as limitações das métricas possam ser repassadas ao modelo TWA serão necessárias restrições adicionais, sendo estas apresentadas na seção 5.3.3.

5.3.1 Congestionamento e Tráfego Retransmitido

Conforme os trabalhos que tratam do VTD, as instâncias são determinadas pelo número total de transceptores utilizados, que é uma função do grau lógico da rede. Pois o número total de transceptores utilizados é dado justamente pelo número de nós da rede multiplicado pelo seu grau lógico (RAMASWAMI; SIVARAJAN; SASAKI, 2009).

Deste modo, para cada grau lógico geramos aleatoriamente uma amostra de soluções para a topologia lógica da rede. Modelos de programação linear foram implementados em *AMPL* (FOURER; GAY; KERNIGHAN, 1990) para distribuir o tráfego da rede sobre cada topologia lógica, através do solver *GLPK* (*GNU Linear Programming Kit* - gnu.org/software/glpk). Obtendo, para cada topologia lógica gerada, a distribuição de tráfego ótima para o congestionamento da rede e também para o total do tráfego retransmitido (ALMEIDA et al., 2006).

Dessas amostras estimamos a média e o desvio padrão das populações para o congestionamento e para o tráfego retransmitido, garantindo a representatividade de cada amostra. Esta estimativa se classifica como um problema de Estatística Indutiva (NETO, 2002), que tem como objetivo tirar conclusões probabilísticas sobre aspectos das populações, com base na observação de amostras extraídas das mesmas. Estas amostras servem de base para inferências que são feitas a cerca das respectivas populações. O processo de estimação de parâmetros utilizado é denominado de estimação por intervalo, onde é contruído um intervalo, o qual deve, com probabilidade conhecida, conter o parâmetro.

Chamamos de estimador a quantidade, calculada em função dos elementos da amostra, que será usada no processo de estimação do parâmetro desejado. O estimador é uma estatística e será, portanto, uma variável aleatória caracterizada por uma distribuição de probabilidade e seus respectivos parâmetros próprios. E chamaremos de estimativa cada particular valor assumido por um estimador.

Ao intervalo que, com probabilidade conhecida, deverá conter o valor real do parâmetro chamaremos intervalo de confiança para esse parâmetro. À probabilidade, que será designada por $1 - \alpha$, de que um intervalo de confiança contenha o valor do parâmetro chamaremos grau de confiança do respectivo intervalo. Dessa forma, vemos que α será a probabilidade de erro na estimação por intervalo, isto é, a probabilidade de errarmos ao afirmar que o valor do parâmetro está contido no intervalo de confiança. Será suposto que os intervalos de confiança são simétricos em probabilidade, isto é, a probabilidade do parâmetro ficar fora do intervalo à sua esquerda é igual à probabilidade de ficar fora à direita, ambas iguais a $\alpha/2$. A partir do exposto, pode-se afirmar que o intervalo conterá ou não o parâmetro, com probabilidade $1 - \alpha$ e α , respectivamente.

Para determinar o tamanho necessário de cada amostra, de modo que estas sejam representativas para cada parâmetro estatístico estimado, utilizamos o método iterativo apontado por (NETO, 2002). Para a média, o intervalo de confiança pode ser obtido por:

$$e_0 = Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Onde $e = 2 \cdot e_0$ é a margem de erro, a amplitude de intervalo de confiança, Z_p é a distribuição normal reduzida, $1 - \alpha$ é o nível de confiança exigido, σ é o desvio padrão populacional e n é o tamanho da amostra. Mas ao invés de estimar o intervalo de confiança, desejamos, a partir de uma margem de erro e um nível de confiança desejados, estimar o tamanho de amostra necessário. Resolvendo a equação acima para n , temos:

$$n = \left(\frac{Z_{\alpha/2} \cdot \sigma}{e_0} \right)^2$$

Onde n agora é a estimativa para o tamanho da amostra. Embora não conheçamos o desvio padrão populacional σ , a relação a seguir permite a substituição pela sua estimativa amostral s :

$$t_{(n-1, \alpha/2)} \cdot s = Z_{\alpha/2} \cdot \sigma$$

Onde $t_{(n-1, p)}$ é a distribuição t de *Student* com $n - 1$ graus de liberdade e probabilidade p , que faz compensar o efeito da maior incerteza embutida pelo uso de s no lugar de σ . Assim, ficamos então com:

$$n = \left(\frac{t_{(n'-1, \alpha/2)} \cdot s}{e_0} \right)^2 \quad (5.3.1)$$

Onde n' é o tamanho da amostra utilizada para calcular s . Agora, se $n \leq n'$, a amostra já terá sido suficiente para a estimação da média. Caso contrário, mais elementos devem ser retirados aleatoriamente da população e adicionados à amostra, até que esta forneça uma estimativa $n \leq n'$.

Procedemos de forma análoga para estimar o tamanho de amostra representativa para a estimativa do desvio padrão, cujo intervalo de confiança pode ser obtido por (NETO, 2002):

$$e' = \sqrt{\frac{(n' - 1) \cdot s^2}{\chi_{(n'-1, 1-\alpha/2)}^2}} - \sqrt{\frac{(n' - 1) \cdot s^2}{\chi_{(n'-1, \alpha/2)}^2}} \quad (5.3.2)$$

Onde $\chi^2_{(n-1,p)}$ é a distribuição *Qui-Quadrado* com $n - 1$ graus de liberdade e probabilidade p e e' é a amplitude do intervalo de confiança dada pela amostra. Neste caso, vamos incrementando o tamanho da amostra (n') até que $e' \leq e$, ou seja, até que o erro calculado seja menor ou igual ao exigido.

Em ambas as estratégias, para a média e para o desvio padrão, é certo que encontraremos um valor em que n' seja representativo, pois s converge para σ quando n' aumenta. Todavia, quanto maior for o nível de certeza exigido e quanto menor for o erro aceito, maior será o tamanho de amostra requerido.

Para definição das formulações apresentadas para estimação da média, é suposto que a distribuição da média amostral é normal (NETO, 2002). Isso ocorrerá se a população for normalmente distribuída ou, caso contrário, com boa aproximação, se a amostra for grande. Em nosso caso de estudo, não se conhece de antemão a distribuição da população, porém o Teorema Central do Limite garante que se a amostra for suficientemente grande, no caso de população infinita ou amostragem com reposição, a distribuição da média amostral será aproximadamente normal. E ainda, quanto maior n' mais a distribuição da média amostral se aproximará da normal. Para verificar se n' é grande o suficiente seria necessário gerar várias amostras e verificar se a distribuição das médias se aproxima da normal. Para não ser necessário realizar essa verificação para todas as instâncias, foram feitos testes pilotos que estimam um tamanho mínimo para n' , de forma que a distribuição das médias amostrais seja aproximadamente normal.

No caso da formulação para estimação do desvio padrão, a equação 5.3.2 é exata para o caso de população normalmente distribuída (NETO, 2002). No entanto, em nossa aplicação não podemos fazer esta afirmativa de antemão, sendo aplicado de forma aproximada este método de estimação. Conforme será apresentado no capítulo 6, a partir dos tamanhos de amostras calculados utilizando as equações 5.3.1 e 5.3.2, e ainda, considerando observações experimentais, a convergência para as estimativas da média e do desvio padrão foi alcançada satisfatoriamente.

Para calcular as distribuições t de *Student* e *Qui-Quadrado* foi utilizado o *Math Toolkit*, uma biblioteca da linguagem de programação C++ do repositório *Boost* (KARLSSON, 2005).

Os parâmetros estimados para o congestionamento darão suporte a escolha da restrição de capacidade do TWA. E por sua vez os parâmetros estimados para o tráfego retransmitido serão utilizados para definir um limite superior para essa métrica.

5.3.2 Custo Físico

Neste trabalho, a rede é projetada desde sua topologia física. A metodologia proposta visa obter uma alocação de recursos suficiente para atender aos requisitos impostos da forma mais econômica possível, mas sem sacrificar o desempenho da rede. Mas, pela própria estrutura do modelo TWA, decisões de menor complexidade são deixadas de fora, pois podem ser resolvidas em outras fases do projeto (LIMA, 2010). Por exemplo, não consideramos aqui o dimensionamento dos nós da rede. Pois a estratégia é que uma boa solução nesta fase do projeto facilite as fases posteriores.

Outro fator a ser considerado é que a solução para a topologia física de uma rede não pode ser muito dependente da instância. Portanto é necessário um critério de qualidade para a topologia física que não considere o número de transceptores, pois estes definirão as instâncias. Como critério de qualidade, propomos aqui garantir que as redes criadas não sejam mais onerosas, proporcionalmente, do que alguma rede real bem conhecida.

Para estabelecer essa proporcionalidade, devemos obter uma razão adimensional que represente cada rede física, para então confrontá-las. Os parâmetros e variáveis utilizados a seguir foram definidos no capítulo anterior.

Isso pode ser feito tomando a razão entre a média por nó do custo de ligações físicas existentes ($FE = (1/N) \cdot \sum_{mn} C_{mn} \cdot D_{mn}$) e a média do custo de todas as ligações físicas possíveis ($FP = \sum_{mn} C_{mn} / (N^2 - N)$). O que resulta na definição 1.

Definição 1 (TIF). A Taxa de Instalação Física (TIF) é dada pela relação:

$$TIF = \frac{FE}{FP} = \frac{N \cdot (N-1)}{N} \cdot \frac{\sum_{mn} C_{mn} \cdot D_{mn}}{\sum_{mn} C_{mn}} = (N-1) \cdot \frac{\sum_{mn} C_{mn} \cdot D_{mn}}{\sum_{mn} C_{mn}}$$

Para garantir a proporcionalidade, inicialmente calcula-se a TIF para uma rede real que será usada como referência. E as redes a serem projetadas usam esta TIF de referência como limitante superior. Ou ainda, pode-se embutir um *TIFgap*, exigindo que a rede projetada tenha TIF 20% ou 30% abaixo da TIF de referência, por exemplo. Todavia, exigir um *TIFgap* muito grande pode acarretar em inviabilidade.

5.3.3 Restrições Adicionadas ao TWA

Para incluir limitações superiores (*Upper Bounds*) em variáveis de interesse, definimos a seguir restrições que devem ser adicionadas ao modelo TWA do capítulo 4.

Dados 2. *Upper bounds:*

1. $TT = \text{Número total de transceptores na rede.}$
2. $TR = \text{Total de tráfego retransmitido.}$
3. $TIF = \text{Taxa de Instalação Física.}$

$$\sum_{mnw} B_{mnw}^{mn} \leq TT \quad (5.3.3)$$

$$\sum_{si,jw} T \cdot q_{sw}^{ij} \cdot A_s \leq TR, \text{ com } i \neq s. \quad (5.3.4)$$

$$\sum_{mn} C_{mn} \cdot D_{mn} \leq TIF \cdot \sum_{mn} C_{mn} / (N - 1), \text{ com } m < n. \quad (5.3.5)$$

A equação 5.3.3 limita superiormente o número total de transceptores na rede, através da soma de todas as componentes topológicas cujo nó de início (n) é o nó de origem das ligações lógicas. Dessa forma, não há exigência quanto a uniformidade do grau lógico na rede, e podemos obter soluções que demandem de uma quantidade menor de transceptores.

Por sua vez, a equação 5.3.4 limita superiormente a soma de todo o tráfego retransmitido na rede, pela soma de todas as componentes de tráfego cujo nó de origem das ligações lógicas (i) é diferente da fonte de tráfego s . Esse *upper bound* para o tráfego retransmitido nos permite garantir uma qualidade aceitável para a retransmissão do tráfego na rede.

Na restrição 5.3.5, que limita superiormente o custo de instalação da rede física, entra em cena a TIF que será imposta às instâncias da rede.

6 *Resultados*

Com o objetivo de demonstrar experimentalmente o funcionamento da metodologia proposta para o projeto de uma WRON em malhas hierárquicas, foram realizados testes computacionais. Serão apresentados os detalhes dos testes realizados com uma rede de 30 nós. As matrizes de tráfego e distâncias, dados de entrada do problema, foram geradas de forma aleatória, com entradas uniformemente distribuídas. A matriz de tráfego possui valores entre 0 e 100. A matriz de distâncias, obedecendo a desigualdade triangular (CORMEN et al., 2002), teve o posicionamento dos pontos sorteado sobre o quadrante do plano cartesiano com ordenadas e abscissas também entre 0 e 100.

O computador onde foram executados os experimentos possui a seguinte configuração: *Intel Pentium®4 CPU 3.00GHz* de 2 núcleos, com 2048KB de *cache*, 1.5GB de RAM, executando o sistema operacional *Kubuntu 9.04*. Na resolução dos modelos de programação linear e programação linear inteira mista, citados ao longo deste capítulo, a precisão nos cálculos adotada foi de 10^{-6} .

Para facilitar o entendimento das etapas do procedimento, elas serão apresentadas e discutidas em sequência, demonstrando a aplicação da metodologia apresentada no capítulo anterior. A primeira mostra como os nós da rede foram divididos nas sub-redes em hierarquia a partir do método heurístico desenvolvido. Na segunda etapa, são apresentados os resultados das estatísticas de congestionamento e tráfego retransmitido para cada sub-rede, além do cálculo da TIF de referência. Por fim, na última etapa são mostrados os resultados da resolução do modelo TWA para todas as sub-redes.

6.1 **Composição das Sub-Redes**

Para a escolha da estrutura hierárquica da rede, que consiste da separação dos nós em *clusters*, foi utilizado o algoritmo genético descrito na Seção 5.1. Definimos $M_i = 4$ e $M_a = 10$, os tamanhos mínimo e máximo dos *clusters*, respectivamente. A partir daí temos, $Q_i = 3$ e $Q_a = 7$, as quantidades mínima e máxima de *clusters* permitidas na solução para a hierarquia.

Dentro do genético existem pesos que podem ser atribuídos às distâncias dentro e fora do *backbone* e para o tráfego de acesso. Esses pesos dependem do caso concreto em que se deseja analisar. Nestes testes esses pesos foram todos deixados como 1. Contudo, existe um parâmetro de calibragem que deve ser definido para deixar tráfego e distâncias com a mesma ordem de grandeza na função objetivo. Pois nela, o tráfego entra como uma soma simples e as distâncias como uma média, além disso, ambos são calculados sobre valores da mesma ordem de grandeza, valores entre 0 e 100.

Para estabelecer o parâmetro de calibragem (C_a), foi feito um teste piloto gerando aleatoriamente indivíduos para o genético, com redes de diversos tamanhos, variando entre 10 e 100 nós. Empiricamente, estimamos que a esperança para o tráfego de acesso das populações do genético (E_t) é razoavelmente interpolada pela curva:

$$E_t(n) = \frac{2}{3} \cdot A_t \cdot (n^2 - n)$$

Onde $A_t = 100$ e equivale à amplitude do intervalo de onde foram retiradas as entradas da matriz de demandas, neste caso 100. Além disso, estimamos que a esperança da distância média das populações do genético (E_d) é satisfatoriamente aproximada por $E_d = A_d/3$. Onde $A_d = 100$ pode ser visto como a média geométrica entre as amplitudes dos intervalos dos eixos cartesianos aonde foram posicionados os pontos da matriz de distâncias. Como $A_t = A_d$, o parâmetro de calibragem fica definido por:

$$C_a = \frac{E_t}{E_d} = 2 \cdot (n^2 - n)$$

Assim, na rede de 30 em estudo, o fator de calibragem adotado foi $C_a = 1740$. Mas vale ressaltar que estas estimativas devem ser revistas para redes com mais que 100 nós, ou no caso de matrizes de demandas ou distâncias com características diferentes das exploradas neste trabalho.

Para os operadores genéticos da *Galib*, as probabilidades de mutação, crossover e substituição foram mantidas em seus valores *default*. O tamanho da população e o número de populações paralelas foram definidos como 100. A estagnação máxima adotada foi 500, ou seja, o genético é interrompido se ficar 500 gerações sem haver melhora.

O tempo de execução do genético foi de 30.76 minutos. A estagnação definida foi alcançada na geração 5473, tendo o genético visitado mais de 13.7 milhões de soluções. A figura 6.1 mostra a evolução do processo de busca. Como pode ser visto, até por volta da geração número 2400 obteve-se uma evolução acentuada para a solução global, a partir daí os avanços foram

menos significativos. Destaca-se que a variação no valor da função objetivo do genético foi superior a 50%, evidenciando a variabilidade das soluções possíveis e a relevância de um bom método para a escolha da estrutura hierárquica da rede.

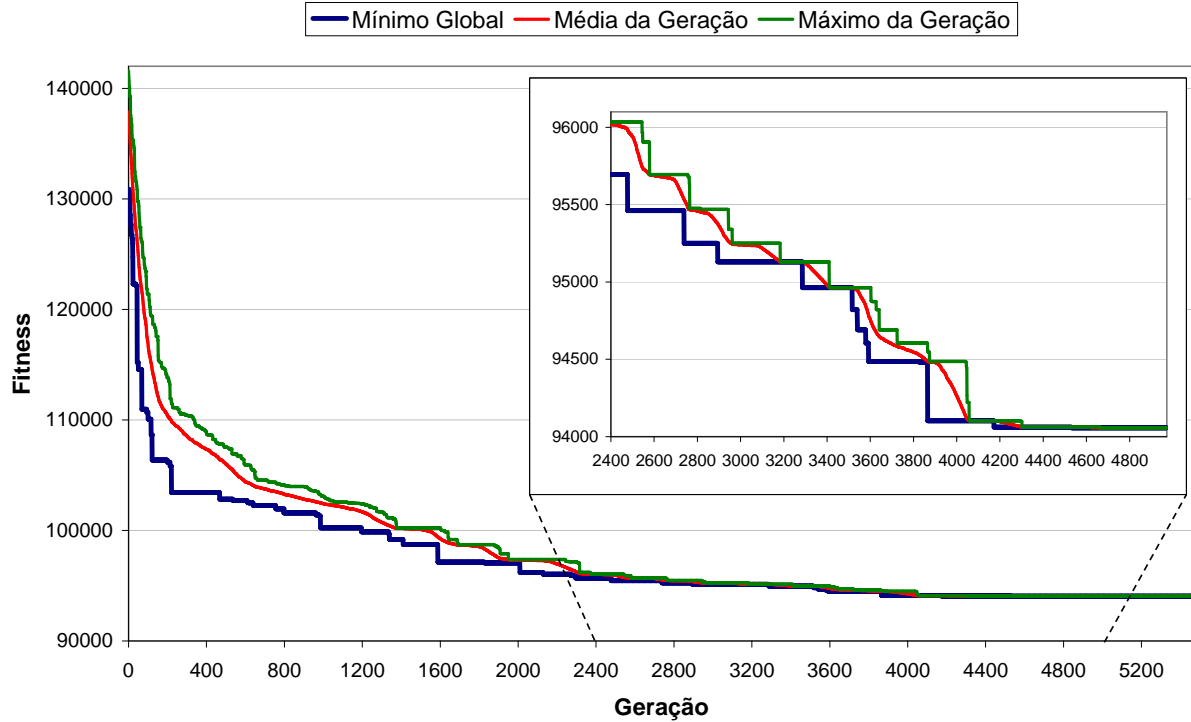


Figura 6.1: Evolução do processo de busca do genético.

A hierarquia obtida, que representa a melhor solução visitada pelo genético, está representada na tabela 6.1. A hierarquia escolhida para a rede de 30 nós é composta de 5 *clusters*, de tamanho variando entre 4 e 8 nós.

Tamanhos:	5	6	7	8	4				
<i>Clusters:</i>									
1 -	7	9	16	26	27				
2 -	28	11	24	25	4	20			
3 -	10	19	3	22	30	17	18		
4 -	15	1	14	23	12	21	29	6	
5 -	13	5	8	2					

Tabela 6.1: Estrutura hierárquica para a rede de 30 nós. Em negrito, os nós do *backbone*.

Para a solução escolhida, a soma das distâncias, antes de calcular a média, para o *backbone* foi $TDist_{bk} = 276.91$, e para os *clusters* foi $TDist_{cl} = 1920.23$. Os pesos para a distância dentro e fora do *backbone* foram definidos como iguais a 1, de forma que no cálculo da média das distâncias o melhor indivíduo possui $M = 24.4$. O total de tráfego de acesso ao *backbone* foi $T =$

$TrafAcesso = 51579.5$, sendo que não foi utilizado peso para o tráfego. Assim, considerando os pesos unitários e o fator de calibragem C_a , o *fitness* do melhor indivíduo foi 94057.6.

As figuras 6.2 e 6.3 mostram os histogramas das soluções visitadas pelo genético durante a sua execução. A figura 6.2, mostra a variedade da distância média entre os nós do *backbone* e dos *clusters*, considerando as possíveis configurações de hierarquia para a rede. Fica evidente que o melhor indivíduo retornado pelo algoritmo, com $M = 24.4$, possui uma excelente configuração com relação ao custo físico da rede.

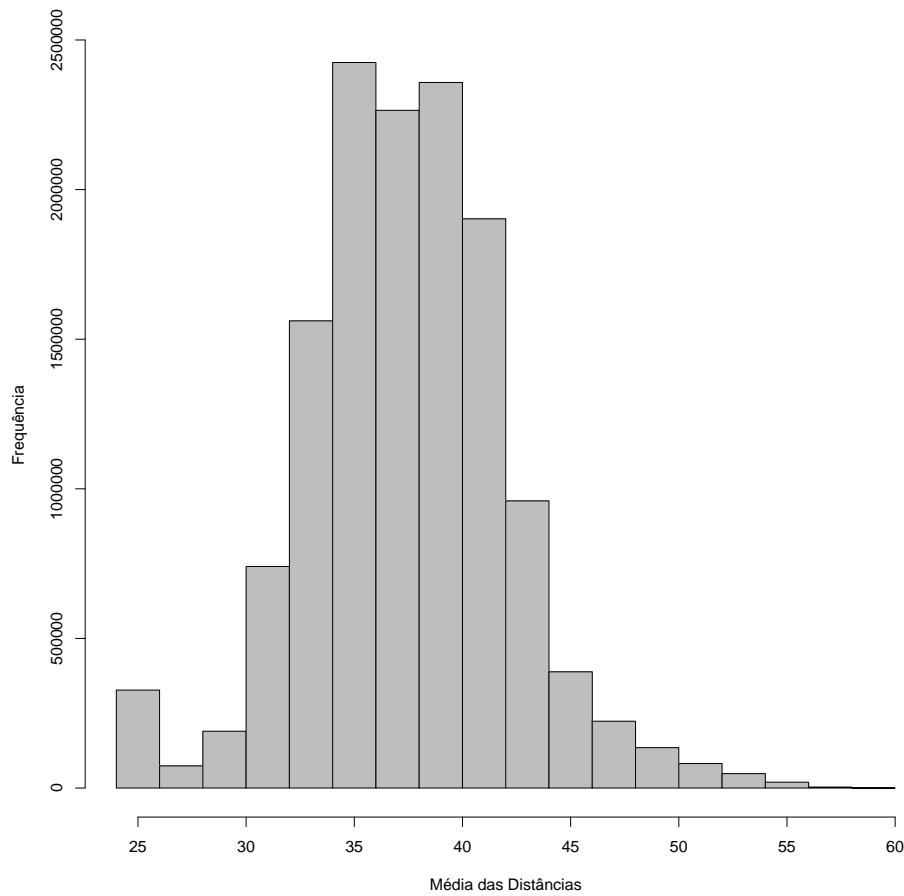


Figura 6.2: Histograma da distância média relativa às soluções visitadas pelo genético.

A figura 6.3 ilustra a variação do tráfego de acesso entre as diversas configurações possíveis para a rede em hierarquia. Como pode ser visto, a solução global retornada pelo genético, com $T = 51579.5$, está situada entre os indivíduos que representam as melhores soluções visitadas para o tráfego de acesso. Em comparação com os resultados observados para distância média, o tráfego de acesso possui uma variação menor com as modificações na configuração da rede. Para a distância a diferença entre a pior e a melhor solução visitada foi superior a 145%, enquanto para o tráfego de acesso essa diferença foi menor que 30%.

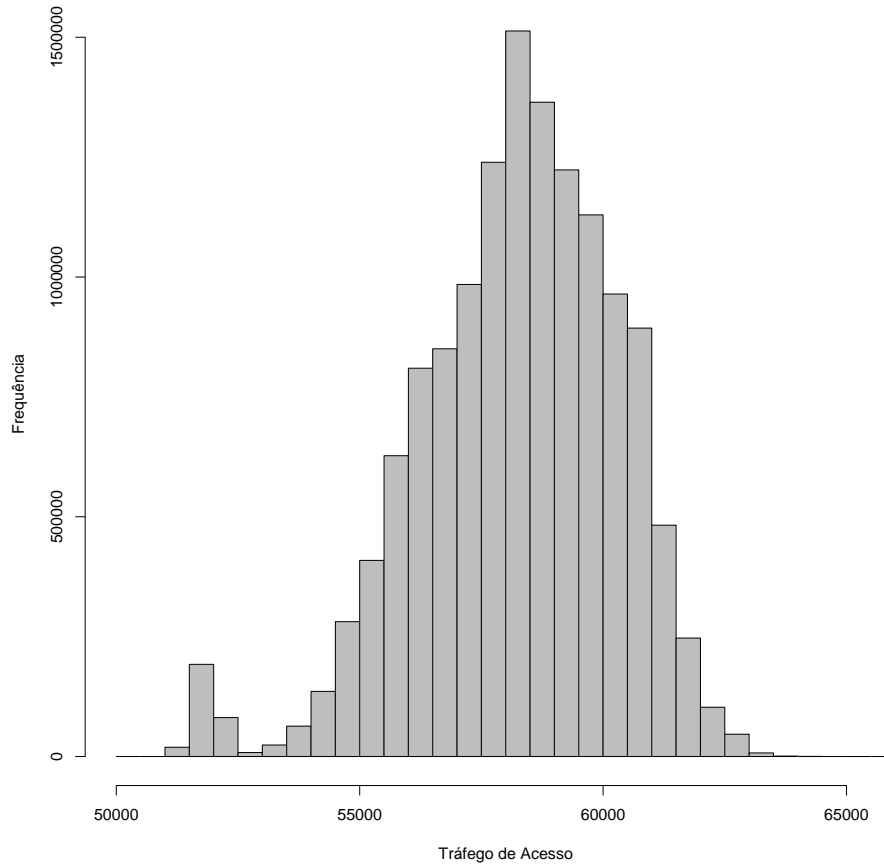


Figura 6.3: Histograma do tráfego de acesso relativo às soluções visitadas pelo genético.

De acordo com o critério de projeto, o posicionamento relativo da distância e do tráfego entre as possíveis estruturas hierárquicas, poderia ser modificada a partir de uma alteração na ponderação entre tráfego de acesso e as distâncias no cálculo da função objetivo.

6.2 Estatísticas de Congestionamento e Tráfego Retransmitido

Para cada sub-rede, *clusters* e *backbone*, as instâncias a serem analisadas são definidas pelo grau lógico. Para uma sub-rede k de tamanho t_k , os graus lógicos considerados são $1, 2, \dots, \lfloor t_k/2 \rfloor$. Para cada instância foram estimados a média e o desvio padrão, ambos com confiança de 99% e margem de erro de 1%. A partir destes parâmetros, o tamanho adotado para cada amostra foi calculado utilizando a formulação apresentada no capítulo 5.

A partir de testes piloto, estimou-se que as estatísticas atingiam estabilidade a partir de 10^3

amostras, para a confiança e o erro adotados. Portanto, esse valor foi utilizado como tamanho mínimo para as amostras.

Para cada grau lógico geramos aleatoriamente uma amostra de soluções para a topologia lógica da rede. Dois modelos de programação linear foram implementados em *AMPL* (FOURER; GAY; KERNIGHAN, 1990) para distribuir o tráfego da rede sobre cada topologia lógica, através do solver *GLPK* (*GNU Linear Programming Kit* - gnu.org/software/glpk). Obtendo, para cada topologia lógica gerada, a distribuição de tráfego ótima para o congestionamento da rede e também para o total do tráfego retransmitido (ALMEIDA et al., 2006). Dessas amostras estimamos a média e o desvio padrão das populações para o congestionamento e para o tráfego retransmitido, garantindo a representatividade de cada amostra.

A figura 6.4 ilustra a convergência da média e desvio padrão do congestionamento para um caso específico, *cluster* 4 com grau lógico 1. Pode-se notar que neste caso, para ambas as estatísticas, 10^3 amostras foram suficientes para atingir razoável estabilidade. Mais adiante será mostrado que tanto para o congestionamento quanto para o tráfego retransmitido, no caso de grau lógico 1, esse tamanho de amostra foi suficiente para todas as sub-redes, segundo o método de estimação adotado. Na figura 6.4 as linhas vermelhas ilustram o valor da média e do desvio padrão referentes ao tamanho de amostra estimado.

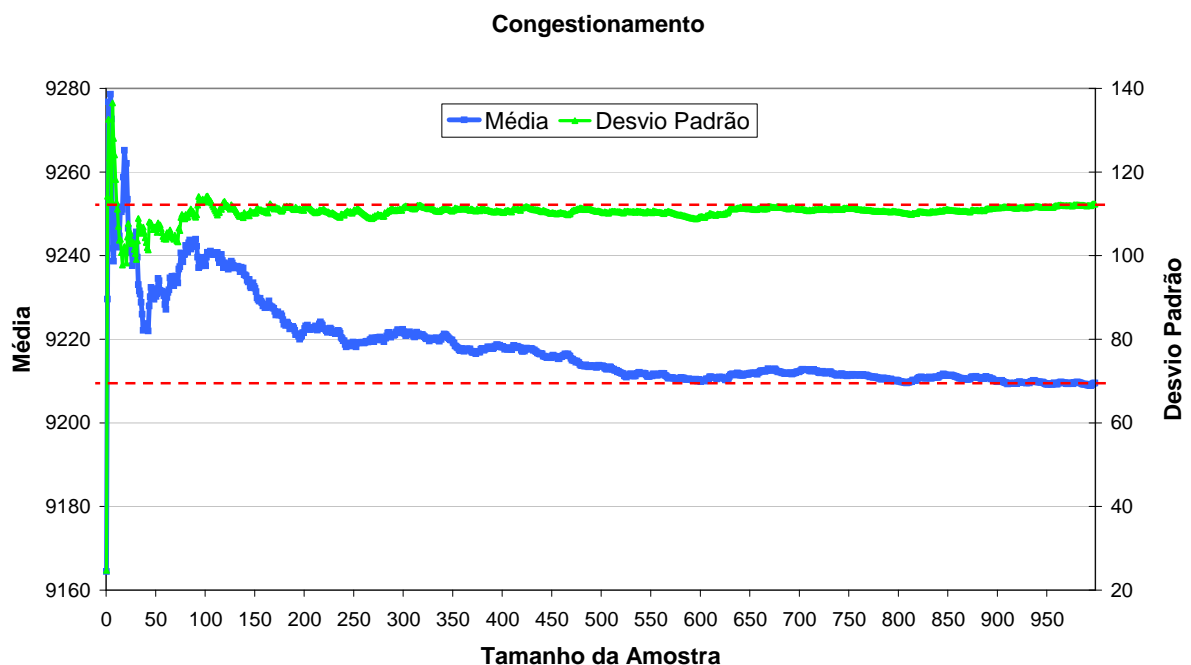


Figura 6.4: Convergência da média e desvio padrão em função do tamanho da amostra.

No decorrer deste capítulo, quando se fizer referência às sub-redes em estudo, a codificação que está apresentada na tabela 6.2 será adotada para identificação do *backbone* e dos *clusters*.

Sub-rede	Tamanho	Identificação (ID)
<i>Backbone</i>	5	bk.t5
<i>Cluster 1</i>	5	cl1.t5
<i>Cluster 2</i>	6	cl2.t6
<i>Cluster 3</i>	7	cl3.t7
<i>Cluster 4</i>	8	cl4.t8
<i>Cluster 5</i>	4	cl5.t4

Tabela 6.2: Identificação das sub-redes.

A tabela 6.3 apresenta as legendas para as tabelas 6.4, 6.5 e 6.6, que serão utilizadas na apresentação dos resultados obtidos para as estatísticas de congestionamento e tráfego retransmitido nas sub-redes. Cada linha dessas tabelas representa uma instância do processo de estimação, sendo definida por uma sub-rede e um grau lógico. O tempo total para o processo de estimação de parâmetros estatísticos foi de 50 minutos.

GI	=	Grau lógico
ID	=	Identificação da sub-rede
LB	=	<i>Lower bound</i>
HLDA	=	Heurística HLDA
Min	=	Mínimo amostral
Max	=	Máximo amostral
MED	=	Média amostral
DP	=	Desvio padrão amostral
MEDn	=	Tamanho da amostra para média
DPn	=	Tamanho da amostra para desvio padrão
t(s)	=	Tempo de execução em segundos
CongDs	=	Desvios padrão para o UB de congestionamento
UBCong	=	<i>Upper bound</i> para congestionamento
TRetrDs	=	Desvios padrão para o UB de tráfego retransmitido
UBTRetr	=	<i>Upper bound</i> para tráfego retransmitido

Tabela 6.3: Legenda para as tabelas 6.4, 6.5 e 6.6

Nas tabelas 6.4 e 6.5, a coluna LB representa um *lower bound* para congestionamento e tráfego retransmitido, respectivamente. As cinco colunas que a seguem, são apresentadas em valores percentuais relativamente ao *lower bound*, mais especificamente, quantos por cento acima do mesmo, com exceção do desvio padrão, coluna *DP*, que é uma relação percentual direta. Como *lower bound* para o congestionamento foi utilizado o MTB (*Minimun Traffic Bound*), enquanto para o tráfego retransmitido utilizou-se o FTB (*Forwarded Traffic Bound*). Ambos os métodos para cálculo de *lower bounds*, foram apresentados em (LIMA, 2010). Nessas mesmas tabelas, para efeito de comparação, a coluna HLDA apresenta respectivamente os

resultados de congestionamento e tráfego retransmitido obtidos utilizando a conhecida heurística HLDA (RAMASWAMI; SIVARAJAN, 1996).

Gl	ID	LB	HLDA	Min	Max	MED	DP	MEDn	DPn	t(s)
1	c5.t4	4196	3%	3%	4%	3%	0.3%	1000	1000	1
1	bk.t5	8897	98%	97%	110%	100%	3.1%	1000	1000	3
1	c1.t5	5563	7%	5%	8%	7%	0.8%	1000	1000	3
1	c2.t6	5928	17%	10%	18%	13%	2.6%	1000	1000	8
1	c3.t7	7264	10%	7%	14%	10%	1.4%	1000	1000	19
1	c4.t8	8150	13%	10%	17%	13%	1.4%	1000	1000	42
2	c5.t4	2098	0%	0%	110%	35%	46.0%	13912	6927	33
2	bk.t5	4448	59%	21%	230%	71%	52.0%	17924	8888	89
2	c1.t5	2782	4%	0%	110%	16%	35.0%	8262	4123	38
2	c2.t6	2964	0%	0%	120%	13%	31.0%	6457	3450	67
2	c3.t7	3632	100%	0%	120%	9%	25.0%	3994	2170	86
2	c4.t8	4075	0%	0%	120%	8%	24.0%	3655	1895	186
3	c2.t6	1976	16%	0%	220%	24%	46.0%	13759	6606	148
3	c3.t7	2421	200%	0%	210%	14%	35.0%	7830	3857	174
3	c4.t8	2717	0%	0%	210%	13%	34.0%	7905	3773	416
4	c4.t8	2037	0%	0%	320%	18%	34.0%	7587	3712	404

Tabela 6.4: Resultado das estatísticas para o congestionamento.

Gl	ID	LB	HLDA	Min	Max	MED	DP	MEDn	DPn	t(s)
1	c5.t4	4500	92%	92%	94%	93%	1.1%	1000	1000	1
1	bk.t5	44000	16%	15%	22%	18%	2.3%	1000	1000	3
1	c1.t5	8500	100%	95%	110%	100%	4.2%	1000	1000	3
1	c2.t6	13000	110%	88%	110%	99%	5.6%	1000	1000	8
1	c3.t7	19000	95%	82%	100%	94%	5.7%	1000	1000	19
1	c4.t8	27000	100%	90%	110%	100%	4.1%	1000	1000	42
2	c5.t4	1400	100%	89%	290%	140%	66.0%	28994	14406	69
2	bk.t5	14000	58%	17%	100%	37%	21.0%	2964	1495	15
2	c1.t5	2700	170%	93%	330%	140%	57.0%	21766	11025	100
2	c2.t6	3900	120%	81%	320%	150%	45.0%	13389	6648	140
2	c3.t7	5300	270%	82%	350%	160%	40.0%	10736	5335	230
2	c4.t8	7700	150%	91%	370%	150%	35.0%	7946	3855	400
3	c2.t6	2400	150%	86%	460%	150%	56.0%	20890	10186	230
3	c3.t7	3700	360%	84%	410%	130%	40.0%	10798	5450	240
3	c4.t8	5000	98%	91%	400%	140%	36.0%	8729	4238	460
4	c4.t8	3500	110%	97%	530%	140%	38.0%	9595	4647	510

Tabela 6.5: Resultado das estatísticas para o tráfego retransmitido.

Uma evidência da qualidade das amostras de topologia visitadas para estimação das estatísticas, é a comparação dos mínimos amostrais com a heurística HLDA, em todos os casos a melhor topologia sorteada obteve resultado melhor ou igual ao HLDA. Outro ponto de destaque

é que para o congestionamento, o *lower bound* calculado possui excelente qualidade visto que é muito próximo dos mínimos amostrais. Quanto ao tráfego retransmitido, o mesmo não se observou, a não ser para o *backbone* onde os mínimos amostrais se aproximaram do *lower bound*.

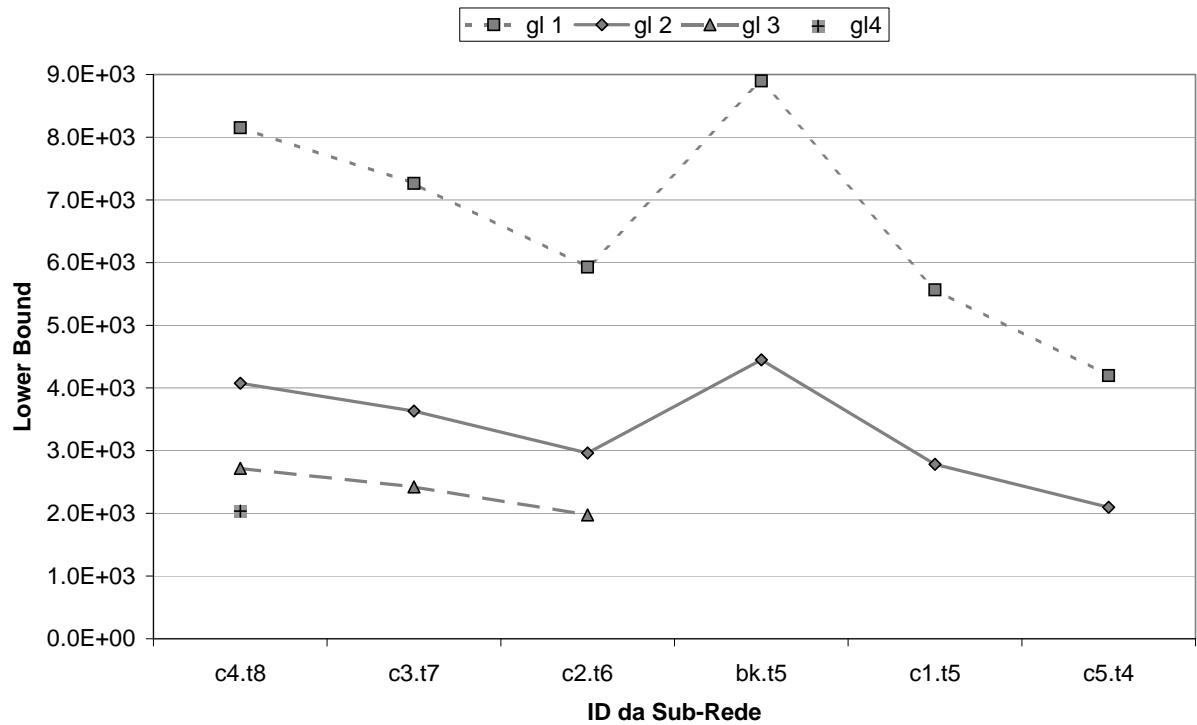


Figura 6.5: *Lower bounds* para o congestionamento.

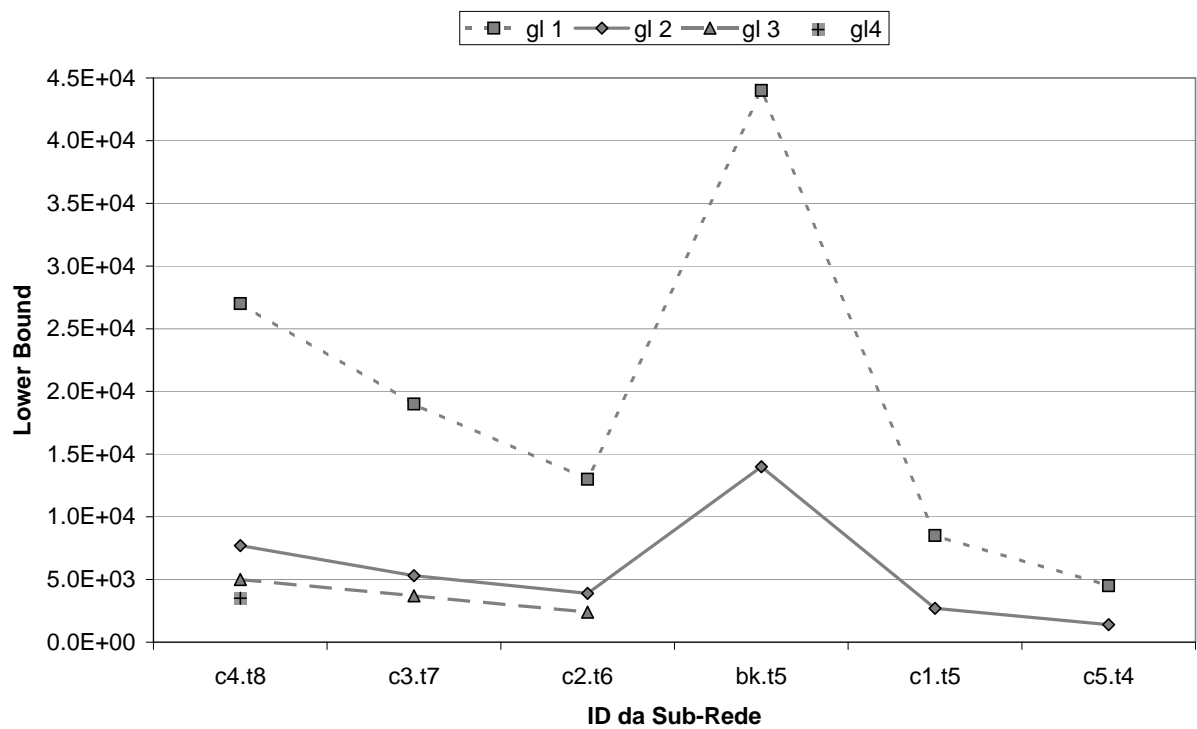


Figura 6.6: *Lower bounds* para o tráfego retransmitido.

Nas tabelas 6.4 e 6.5, observa-se casos onde a grande variabilidade das topologias visitadas leva a desvios padrão superiores a 50%. Para o congestionamento, chamam a atenção as instâncias que possuem desvio padrão superior a média. Para o caso de grau lógico 1, rede em anel, destaca-se a pequena diferença entre máximo e mínimo. A pouca variabilidade desse tipo de topologia é reforçada pelo pequeno desvio padrão observado nestas instâncias.

As figuras 6.5 e 6.6 apresentam graficamente as variações dos *lower bounds* de congestionamento e tráfego retransmitido, respectivamente. Nelas é possível observar o comportamento dos mesmos em função do grau lógico em cada sub-rede.

Conforme mencionado no capítulo anterior, o resultado principal dessa etapa de estimação de estatísticas para congestionamento e tráfego retransmitido é a definição de *upper bounds* para os mesmos. A partir das estatísticas obtidas pode-se definir uma regra para esses *upper bounds* que represente valores de boa qualidade para essas métricas. Os valores definidos em função das estatísticas serão repassados para o modelo TWA na etapa final do processo de resolução.

Conforme definido nas equações 6.2.1 e 6.2.2, os *upper bounds* foram calculados como a média menos uma quantidade de desvios padrão, tanto para o congestionamento (*CongDs*) quanto para o tráfego retransmitido (*TRetrDs*). Para as duas métricas essa quantidade de desvios padrão é múltipla de 0.5 e foram definidas como a maior possível, desde que o *upper bound* resultante não ultrapasse o respectivo mínimo amostral. A tabela 6.6 mostra os valores resultantes desta regra.

$$UBCong = MED - CongDs \cdot DP \quad (6.2.1)$$

$$UBTRetr = MED - TRetrDs \cdot DP \quad (6.2.2)$$

Para os casos em que $CongDs = 0$, significa que o *upper bound* para o congestionamento foi definido como a própria média amostral (*MED*).

Além dos parâmetros de controle *UBCong* e *UBTRetr*, que obtemos a partir das estatísticas, para garantir que a solução obtida para a topologia física com o modelo TWA não seja muito dependente da instância, como foi comentado na seção 5.3.2, calculamos a *TIF* para a rede NF-SNET (RAMASWAMI; SIVARAJAN; SASAKI, 2009). O valor calculado foi $TIF_{NFSNET} = 1.82$, e este será o valor de referência sobre o qual ainda exigimos uma folga para baixo (*TIFgap*), para dessa forma definir o *upper bound* para o custo físico (*UBCFis*) de cada sub-rede. Conforme será detalhado na próxima seção, o *TIFgap* utilizado variou entre 70% e 20%.

ID	GL	CongDs	UBCong	TRetrDs	UBTRetr
bk.t5	1	1.0	17636	1.0	51437
bk.t5	2	0.5	6434	0.5	18088
c1.t5	1	1.0	5883	1.0	16835
c1.t5	2	0.0	3224	0.5	5774
c2.t6	1	1.0	6558	1.5	24601
c2.t6	2	0.0	3359	1.5	7092
c2.t6	3	0.5	2002	1.0	4569
c3.t7	1	1.5	7836	1.5	35938
c3.t7	2	0.0	3950	2.0	9625
c3.t7	3	0.0	2768	1.0	6872
c4.t8	1	2.0	8984	2.5	51115
c4.t8	2	0.0	4379	1.5	15137
c4.t8	3	0.0	3071	1.0	9981
c4.t8	4	0.5	2056	1.0	7194
c5.t4	1	1.5	4306	1.0	8555
c5.t4	2	0.5	2353	0.5	2991

Tabela 6.6: *Upper bounds* calculados para congestionamento e tráfego retransmitido.

$$UBCFis = \frac{TIF_{NFSNET} \cdot (1 - TIF_{gap}) \cdot \sum_{mn} C_{mn}}{N - 1} \quad (6.2.3)$$

A equação 6.2.3 define a fórmula de cálculo para $UBCFis$. Como se pode notar, partindo de um valor de referência (TIF_{NFSNET}), um fator de qualidade (TIF_{gap}) e a somatória dos custos de instalação de uma sub-rede, neste trabalho representado pela matriz de distâncias, é possível calcular o *upper bound* para cada instância do problema.

6.3 Resolução das Sub-redes com TWA

Como iremos procurar soluções resolvendo um modelo de programação linear inteira, nestes testes, não estamos utilizando a capacidade do TWA de produzir soluções com múltiplas ligações físicas, com o intuito de reduzir o número de variáveis binárias no modelo. Entretanto ainda são admitidas soluções com múltiplas ligações lógicas, pois no TWA isso não interfere no número de variáveis.

Todas as estatísticas de congestionamento e tráfego retransmitido foram calculadas utilizando como instâncias as combinações de sub-rede e grau lógico. Destaca-se que as restrições do modelo TWA utilizadas para controle de grau lógico, na prática não limitam o grau lógico de cada nó da forma tradicional (RAMASWAMI; SIVARAJAN, 1996), efetivamente o que é limitado é o total de transceptores utilizados na topologia virtual, conforme a restrição 5.3.3. Dessa

forma, as instâncias do TWA também são definidas pelo grau lógico, mas como referência, pois o que será limitado é o total de transceptores, que são no máximo o número de nós da sub-rede vezes o grau lógico da instância.

Os valores de *upper bounds* utilizados para tráfego retransmitido e congestionamento são os apresentados na tabela 6.6. Com relação à limitação do custo físico, iniciamos com um *TIFgap* de 40% pois em testes piloto este se mostrou promissor. Reduzimos o *TIFgap* nas situações em que não foi possível encontrar solução viável. Além disso, aumentamos o *TIFgap* enquanto ainda fossem encontradas soluções, visando obter o melhor custo físico possível que ainda atendesse às demais exigências. Em ambos os casos, o valor do *TIFgap* é alterado em 10%, para cima ou para baixo.

Para cada *TIFgap*, a estratégia adotada foi, partindo do menor grau lógico ($Gl = 1$), fixar nos valores mínimos o número de comprimentos de onda, e executar o modelo básico do TWA com as adições indicadas na Seção 5.3.3. Um *solver* para problemas MILP é instanciado com essa configuração. Enquanto o *solver* retornar que o problema é insolúvel (MUKHERJEE, 2006), o número de comprimentos de onda (W) será incrementado, até um limite definido. Atingido esse limite, o *TIFgap* é reduzido de 10% e W retornado para seu valor inicial.

Se o *solver* não encontrar uma solução viável dentro de 10^4 segundos, então ele é abortado e é aplicado o mesmo procedimento de quando a solução é insolúvel. Isso é feito com o objetivo de folgar as restrições de modo que o *solver* possa encontrar uma solução viável mais rapidamente, sacrificando um pouco a qualidade das restrições. Essas situações fazem parte do método e o seu tempo é contabilizado conjuntamente com o das instâncias insolúveis. Um método de uso similar para o modelo TWA foi apresentado anteriormente em (LIMA et al., 2009).

Se uma solução viável é encontrada, o *solver* é interrompido, a solução é registrada e o grau lógico é incrementado, dando continuidade ao processo. Vale lembrar que o *solver* é interrompido pois todos os parâmetros de interesse já estão limitados satisfatoriamente, portanto qualquer solução viável já atende o que foi exigido. Não sendo necessário perseguir otimalidade neste ponto.

Como, nesta modelagem, W está diretamente relacionado a quantidade de variáveis, é mais conveniente começar com $W = 1$. Disso decorre a escolha de também começarmos com $Gl = 1$. A maior prioridade para a incrementação é dada ao Gl , pois variando este temos instâncias diferentes. A menor precedência ficou para W , pois quanto menor ele for menores precisarão ser os custos de instalação da rede.

O pseudo código 1 ilustra o processo iterativo de resolução do modelo TWA para cada

sub-rede.

Pseudo Código 1 (Uso do TWA). Para cada sub-rede k de tamanho t_k , deve-se executar:

Dados de Entrada da sub-rede:

Matrizes de tráfego e distância; $UBCong$; $UBTRetr$; TIF_{NFSNET}

```

1      FOR  $Gl = 1$  TO  $Gl = \lfloor t_k/2 \rfloor$  DO
2           $W = 1$ ;  $TIF_{gap} = 40\%$ ;
3          DO UNTIL  $W = 5$  OR Solução Viável
4               $SOLVER_{[TWA]}$ ;
5               $W = W + 1$ ;
6          END DO
7          IF NOT Solução Viável AND  $10\% \leq TIF_{gap} \leq 40\%$ 
8               $TIF_{gap} = TIF_{gap} - 10\%$ ;
9               $W = 1$ ;
10             GO TO step (3)
11         END IF
12         IF Solução Viável AND  $TIF_{gap} \geq 40\%$ 
13              $TIF_{gap} = TIF_{gap} + 10\%$ ;
14              $W = 1$ ;
15             GO TO step (3)
16         END IF
17     END FOR

```

Utilizamos o *solver* SCIP (*Solving Constraint Integer Programs* - scip.zib.de) para encontrar as soluções viáveis para as diversas instâncias do TWA. Além de calcular congestionamento e tráfego retransmitido para as estatísticas, como foi descrito na seção 5.3.1, o GLPK também foi usado para interpretar o modelo AMPL, gerando a entrada de dados para o SCIP. Vale observar que o SCIP e o GLPK são *softwares* livres, de código fonte aberto e de distribuição gratuita.

A tabela 6.7 apresenta as legendas para a tabela 6.8, que será utilizada na apresentação dos resultados obtidos para o projeto completo das sub-redes utilizando o modelo TWA. Nesta tabela, cada linha representa uma instância, composta por uma sub-rede e um grau lógico.

Conforme descrito na seção anterior, a limitação de custo físico foi definida com base na TIF da rede real NFSNET, como definido na equação 6.2.3. Os valores de $UBCFis$, foram calculados a partir desse valor de referência para a TIF, menos um percentual TIF_{gap} , que representa uma tentativa de forçar uma boa qualidade no resultado do projeto da rede física.

ID	=	Identificação da sub-rede
Gl	=	Grau lógico
W	=	Número de comprimentos de onda
TIFgap	=	Folga em relação à TIF da NFSNET
TTranc	=	Total de transceptores
CFis	=	Custo físico
TIF	=	TIF resultante
Cong	=	Congestionamento
TRetr	=	Tráfego retransmitido
Ti(s)	=	Tempo de execução da instância
Tt(s)	=	Tempo de execução total

Tabela 6.7: Legenda para a tabela 6.8.

Na coluna TIF da tabela 6.8, está apresentado o valor da TIF calculado para a solução obtida para a topologia física. O número de comprimentos de onda (W) necessários para atender às demandas de tráfego, que é um fator relevante do ponto de vista de complexidade e custo de uma rede, variou entre 1 e 4.

ID	Gl	W	TIFgap	TTranc	CFis	TIF	TRetr	Cong	Ti(s)	Tt(s)
bk.t5	1	1	70%	5	51	0.54	51084	17526	0.1	5
bk.t5	2	2	70%	10	51	0.54	17667	6434	0.4	15
c1.t5	1	1	20%	5	79	1.42	16635	5882	1	81842
c1.t5	2	2	20%	9	79	1.42	5755	3003	11	95
c2.t6	1	2	40%	6	68	1.07	24484	6523	98	831
c2.t6	2	3	40%	12	68	1.07	5556	2730	9	225
c2.t6	3	4	40%	16	68	1.07	4267	2002	74	509
c3.t7	1	1	30%	7	85	1.24	35709	7821	37	221110
c3.t7	2	3	30%	13	87	1.26	7755	2961	215	24271
c3.t7	3	3	30%	18	84	1.23	6824	2768	2859	17586
c4.t8	1	1	40%	8	80	1.09	50807	8982	698	41074
c4.t8	2	2	40%	16	79	1.06	7976	3582	17	21350
c4.t8	3	2	40%	16	79	1.07	7069	3050	58	3999
c4.t8	4	3	40%	22	77	1.04	6590	2056	587	23027
c5.t4	1	1	20%	4	65	1.36	8546	4304	0.4	9
c5.t4	2	2	20%	7	65	1.36	233	1515	0.1	6

Tabela 6.8: Resumo dos resultados com TWA.

Os resultados obtidos para TIF das sub-redes mostra que o processo de busca iniciado com o genético, onde os nós foram agrupados por proximidade, gerou uma solução final de boa qualidade, visto que foi possível utilizar fatores de qualidade (*TIFgap*) em todos os casos. Destaca-se o resultado da TIF para o *backbone* (0.54), sendo muito inferior à TIF de referência (1.82). O pior resultado foi obtido para *c1.t5*, onde a TIF obtida foi de 1.42.

Para duas instâncias, *c1.t5* e *c3.t7*, ambas com grau lógico 1, mesmo com um *TIFgap* reduzido a 0.0% não foi possível encontrar solução viável. Então relaxamos os *upper bounds* para congestionamento e tráfego retransmitido, reduzindo em 0.5 o número de desvios padrão abaixo da média para ambos, nas duas instâncias. Os valores mostrados na tabela 6.6 foram apresentados com essas exceções.

Os experimentos mostraram que o custo físico não possui relação evidente com o grau lógico. Isso fica claro pois para cada sub-rede o valor de *TIFgap* para o qual foi possível obter soluções viáveis foi o mesmo, independente do grau lógico. Além disso, pode-se notar por exemplo que entre o *backbone* e o *cluster 1*, ambos com 5 nós, o *TIFgap* teve uma variação de 50%. Essas características cumprem o objetivo proposto na definição da *TIF* de ser um parâmetro independente do grau lógico e do número de nós da rede.

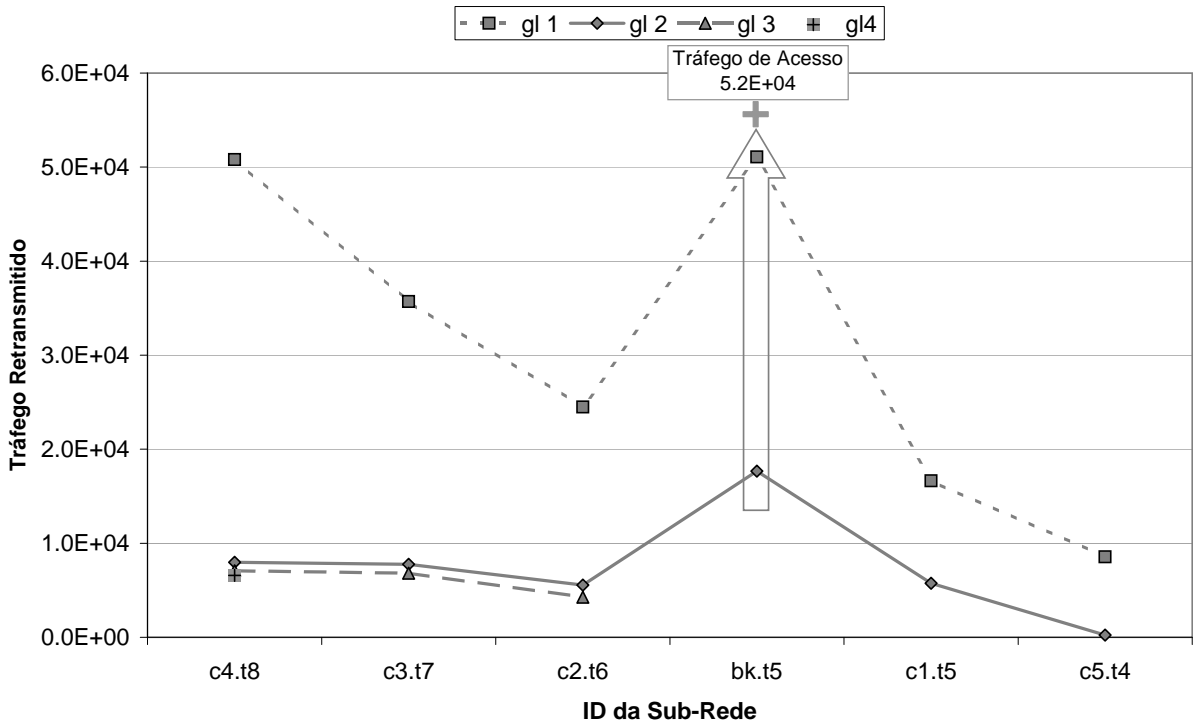


Figura 6.7: Resultados para tráfego retransmitido e de acesso ao *backbone*.

A figura 6.7 mostra o comportamento do tráfego retransmitido em cada sub-rede de acordo com o grau lógico. Além disso, para evidenciar a solução global da rede em hierarquia, ela destaca o tráfego de acesso ao *backbone* que foi calculado na etapa inicial, quando o genético fez a escolha das sub-redes.

Na versão do modelo TWA aqui utilizada, o grau lógico foi representado como uma restrição no número total de transceptores na rede. Como se trata de uma limitação, existem casos onde não são utilizados todos os transceptores permitidos pelo grau lógico definido. Na coluna *TTranc* da tabela 6.8, nota-se que para grau lógico 1 o número de transceptores ficou

igual ao número de nós, configurando anéis, que é a única topologia conexa de grau lógico 1 possível. Fora estes casos, somente para o grau lógico 2, em algumas sub-redes, foi necessário utilizar todos os transceptores permitidos.

As últimas colunas da tabela 6.8, mostram os tempos de execução do processo de busca. Como o modelo TWA foi utilizado de forma iterativa, conforme descrito no pseudo código 1, variando os parâmetros de limitação ou qualidade para as restrições, são apresentados dois tempos para cada instância. A coluna $Tt(s)$ mostra o tempo total, que contempla todas as tentativas de utilização do *solver*, devido às alterações em W e $TIFgap$. Além disso, para os casos de exceção citados acima, onde foi necessário folgar os *upper bounds* de tráfego e congestionamento, também foi contabilizado o tempo das tentativas com os valores originais de desvios padrão abaixo da média ($CongDs$ e $TRetrDs$). Este fato explica os elevados valores de $Tt(s)$ para as instâncias $c1.t5$ e $c3.t7$, com grau lógico 1.

A coluna $Ti(s)$ mostra o tempo relativo apenas à última tentativa de resolução, ou seja, aquela que retornou uma solução viável. Como pode-se notar, na maioria dos casos esse último tempo foi de apenas alguns segundos. O tempo total para resolução de todas as instâncias, somatório da coluna $Tt(s)$, foi de 121 horas. Cabe ressaltar que se a tolerância de tempo para o *solver* retornar uma solução viável fosse reduzida para 10^3 segundos, ao invés de 10^4 , o tempo total para resolução de todas as instâncias seria reduzido consideravelmente, sem se observar uma perda de qualidade relevante. Nesta situação, apenas uma instância teria seu resultado piorado em função da menor tolerância de tempo, seria $c3.t7$ com grau lógico 3, sendo que este foi o único caso em que o tempo $Ti(s)$ para encontrar uma solução viável superou 10^3 segundos.

7 *Conclusões*

Este trabalho explorou o projeto completo de redes ópticas. A metodologia para o projeto de redes apresentado engloba os conhecidos problemas VTD e RWA, além da escolha da topologia física da rede. O objetivo principal do trabalho foi criar uma metodologia que viabiliza o projeto completo de WRON com topologia de malhas hierárquicas, para tanto, aproveitando o potencial da modelagem TWA, juntamente com métodos heurísticos e indução estatística. O procedimento apresentado para a resolução de WRON com topologias hierárquicas, faz uso de um algoritmo genético e da eficiência do modelo TWA, além da estimação de parâmetros de qualidade para as principais métricas que envolvem o projeto da rede.

Uma simplificação adotada foi assumir que as ligações entre os *clusters* e o *backbone* são opacas, ou seja, todos os componentes de uma ligação lógica estão em um mesmo nível da hierarquia. Deste modo, o *backbone* e os *clusters* são vistos como redes independentes, exceto pelo tráfego de acesso, que cruza de um nível hierárquico para outro. Como foi explorado, este tráfego pode ser acumulado nos nós de acesso ao *backbone*, de modo que sua distribuição e roteamento podem ser feitos paralelamente em cada sub-rede. Essas considerações servem de base para permitir que cada sub-rede seja projetada separadamente, eliminando a necessidade de uma modelagem para toda a hierarquia.

As variáveis de interesse consideradas foram o congestionamento da rede, o total do tráfego retransmitido, o custo de instalação da rede física, o número de comprimentos de onda utilizados e o número total de transceptores. Para estabelecer parâmetros de qualidade para congestionamento e tráfego retransmitido, foram estabelecidos critérios de qualidade baseados em estimativas estatísticas das médias e desvios padrões das populações de soluções possíveis, em cada sub-rede. Para o número de comprimentos de onda, foi utilizada uma estratégia iterativa, buscando-se obter soluções com um comprimento de onda apenas e incrementando esse número enquanto não fosse possível encontrar soluções viáveis.

Por sua vez, o custo de instalação da rede física foi controlado garantindo que seja menor, proporcionalmente, que o custo de uma rede real conhecida, aqui considerada a NSFNET (RAMASWAMI; SIVARAJAN; SASAKI, 2009). Considerou-se como custo de instalação a dis-

tância entre os nós da rede, que representa o comprimento dos enlaces de fibra instaladas entre os mesmos. Visando permitir comparação com uma rede de referência, foi definida a taxa de instalação física (*TIF*), que representa um indicador do custo total de instalação da rede física.

Conforme apresentado no capítulo 5, a metodologia proposta é constituída por três etapas sequenciais. Na primeira, são fornecidas as matrizes de demandas de tráfego e distâncias, como dados de entrada do projeto. Um algoritmo genético faz a separação dos nós em sub-redes. Para tanto, utiliza como critério de avaliação o tráfego a ser transferido entre as sub-redes e a distância entre os nós, ou seja, procura-se concentrar o tráfego no interior das sub-redes, e formá-las com nós geograficamente próximos.

Na fase intermediária, são obtidos os limites superiores para o congestionamento e o tráfego retransmitido na rede, por meio de inferência estatística baseada na média e desvio padrão dos resultados das avaliações de amostras de soluções viáveis para o VTD. Estes limites superiores são valores máximos admissíveis para o congestionamento e o tráfego retransmitido, impostos às variáveis do modelo de programação inteira-mista para garantir, com relação a estas métricas, a qualidade da solução a ser obtida na próxima etapa. Nesta mesma etapa, é calculada a *TIF* da rede de referência que serve de base para a imposição de restrições de custo físico às sub-redes.

A última etapa consiste na resolução do modelo TWA para cada sub-rede, respeitando as exigências de qualidade, os limitantes calculados, e visando utilizar o menor número possível de comprimentos de onda e transceptores. O resultado desta última fase fornece a solução do projeto, representado pelas topologias física e lógica, bem como a forma de roteamento e alocação dos comprimentos de onda, em função da distribuição das demandas de tráfego. Esses resultados são referentes às sub-redes em separado, porém, quando agrupados, representam o resultado global para uma WRON com topologia em malhas hierárquicas.

Cabe ressaltar alguns conceitos que serviram de fundamento para a metodologia apresentada. Primeiramente havia o interesse em aproveitar a abrangência e bom desempenho do modelo TWA para o projeto de uma WRON. Considerou-se também como premissa explorar topologias hierárquicas, um tipo de topologia ainda pouco estudado. Conforme já foi amplamente explorado na literatura, o uso de meta-heurísticas na resolução de sub-problemas contribui para a simplificação dos modelos de otimização e redução do custo computacional, por isso a aplicação de um Algoritmo Genético. Além disso, a definição *a priori* de exigências de qualidade para os parâmetros de interesse simplifica a etapa de maior demanda de tempo no procedimento, pois permite utilizar o modelo matemático para obtenção de soluções viáveis de qualidade aceitável, sendo esses critérios de qualidades definidos previamente através de indução estatística e valores de referência, como *lower bounds* e *TIF* de referência.

Esta formulação é mais ampla que as apresentadas na literatura e justamente por este fato existe uma dificuldade de comparação dos resultados aqui apresentados. A partir daí foram utilizados os limites calculados e a *TIF* de uma rede real conhecida para permitir a avaliação da qualidade da solução obtida nos testes computacionais. Os testes realizados para validar experimentalmente a metodologia foram feitos com uma rede de 30 nós.

No capítulo 6 foram apresentados os detalhes referentes ao experimento realizado e resultados observados. O experimento objetivou exemplificar a aplicação da metodologia em uma situação de projeto de rede. O tempo computacional necessário pode ser considerado aceitável se for levado em consideração o porte da rede e a abrangência do projeto que a metodologia permite resolução. Conforme descrito, o tempo de execução poderia ter sido consideravelmente melhorado com uma redução na tolerância de tempo adotada para o modelo TWA retornar uma solução viável na última etapa do procedimento, que constitui-se de um processo iterativo.

Para trabalhos futuros seria relevante submeter a metodologia proposta a uma bateria de testes variados quanto ao tamanho da rede e a característica das matrizes de tráfego e distâncias, analisando matrizes não uniformes. Outras aplicações seriam o projeto de redes com topologia em malha genéricas e as variações permitidas para o modelo TWA, considerando diferentes restrições e funções objetivo. Um outro item que poderia ser melhor explorado é a forma de contabilização do custo físico, empregando modelos mais realistas de avaliação, que incluam os equipamentos ópticos e outras restrições da camada física.

Publicações

Relação da produção bibliográfica do autor desta dissertação.

- **Artigos completos publicados em periódicos**

1. Lima, Marcelo de Oliveira; Lima, F. O.; Oliveira, E. S.; Segatto, M. E. V.. *Um Algoritmo Híbrido para o Planejamento de Redes Ópticas*. REIC - Revista Eletrônica de Iniciação Científica, v. 4, p. 4, 2006.

- **Trabalhos completos publicados em anais de congressos**

1. Lima, M. O.; Lima, F. O.; Segatto, M. E. V.; Almeida, R. T. R.; Oliveira, E. S.. *Projeto Completo de Redes Ópticas em Hierarquia*. In: MOMAG 2010, 14º SBMO - Simpósio Brasileiro de Microondas e Optoeletrônica, 2010.
2. Lima, F. O.; Lima, M. O.; Segatto, M. E. V.; Almeida, R. T. R.. *Projeto Completo de Redes Ópticas*. In: MOMAG 2010, 14º SBMO - Simpósio Brasileiro de Microondas e Optoeletrônica, 2010.
3. Lima, F. O.; Lima, M. O.; Segatto, M. E. V.; Almeida, R. T. R.; Oliveira, E. S.. *Um modelo eficiente para o projeto completo de redes ópticas*. In: XLI SBPO - Simpósio Brasileiro de Pesquisa Operacional, 2009.
4. Silva, M. ; Paiva, M.; Lavagnoli, G.; Lima, M. O.; Segatto, M. E. V.; Oliveira, E. S.; Almeida, R. T. R.. *On Solving HSHR Networks*. In: 6th Conftele - Conference on Telecommunications, 2007.
5. SILVA, M.; Paiva, M.; Lavagnoli, G.; Lima, M. O.; Segatto, M. E. V.; Almeida, R. T. R.; Oliveira, E. S.. *Análise de Redes Ópticas em Anéis Hierárquicos*. In: XXV SBrT - Simpósio Brasileiro de Telecomunicações, 2007.
6. Segatto, M. E. V.; Oliveira, E. S.; Lima, M. O.; Lima, F. O.; Almeida, R. T. R.. *Hybrid approaches for the design of mesh and hierarchical ring optical networks*. In: Proceedings of SPIE06 - Photonics Europe 2006, v. 1.
7. Lima, F. O.; Lima, M. O.; Oliveira, E. S.; Segatto, M. E. V.. *Reformulando o Problema de Projeto de Anéis em Redes Ópticas*. In: Proceedings of 4th ITS - International Information and Telecommunication Technologies Symposium, 2005.

8. Lima, M. O.; Oliveira, E. S.; Pereira, L. C. B.; Almeida, R. T. R.; Segatto, M. E. V.. ***A Hybrid-Combined Algorithm Approach for the Design Topologies and Flow Congestion Minimization of Optical Networks***. In: Proceedings of the 5th Conference on Telecommunications - Conftele, 2005.
9. Lima, M. O.; Maioli, C.; Botelho, T.; Pereira, L. C. B.; Almeida, R. T. R.; Segatto, M. E. V.; Oliveira, E. S.. ***The Design of Hierarchical Self-Healing Rings Networks***. In: Proceedings of the 5th Conference on Telecommunications - Conftele, 2005.
10. Lima, M. O.; Oliveira, E. S.; Segatto, M. E. V.; Almeida, R. T. R.. ***Projeto de Redes Ópticas com Topologia em Anéis Hierárquicos Tolerantes a Falhas***. In: Anais do XXII Simpósio Brasileiro de Telecomunicações - SBrT 05, 2005. p. 1236-1237.
11. Lima, M. O.; Oliveira, E. S.; Pereira, L. C. B.; Almeida, R. T. R.; Segatto, M. E. V.. ***Estratégias com Algoritmos Híbridos para Projeto de Redes Ópticas***. In: SBPO - Simpósio Brasileiro de Pesquisa Operacional, 2004.

Referências Bibliográficas

- ALMEIDA, R. T. et al. Design of virtual topologies for large optical networks through an efficient milp formulation. *Optical Switching and Networking*, v. 3, n. 1, p. 2 – 10, 2006.
- ASSIS, K. D. R.; WALDMAN, H. Topologia virtual e topologia física de redes Ópticas: Uma proposta de projeto integrado. *Revista da Sociedade Brasileira de Telecomunicações*, v. 19, n. 2, p. 119–126, 2004. Rio de Janeiro.
- ASSIS, K. D. R.; WALDMAN, H.; GIOZZA, W. Optical networks: A complete design. *Journal of communication and information systems*, v. 20, p. 81–95, 2005. IEEE ComSoc/SBrT Special Joint Issue.
- BANNISTER, J. A.; FRATTA, L. Topological Design of the Wavelength-Division Optical Network. *INFOCOM*, 1990.
- CORMEN, T. H. et al. *Algoritmos: Teoria e prática*. 2. ed. São Paulo: Elsevier, 2002.
- FOURER, R.; GAY, D. M.; KERNIGHAN, B. W. A Modeling Language for Mathematical Programming. *Management Science* 36 pp. 519-54, 1990.
- GOLDBERG. *Genetic Algorithms in Search, Optimization, and Machine Learning*. [S.l.]: Addison-Wesley Publishing Company, Inc., 2002.
- GUAN, C.; CHAN, V. Efficient Physical Topologies for Regular WDM Networks. *Optical Fiber Communication Conference - OFC*, 1997.
- JAUMARD, B.; MEYER, C.; THIONGANE, B. Comparison of ilp formulations for the rwa problem. *Les Cahiers du GERAD*, v. 66, 2004. ISSN 0711-2440. G-2004-66.
- KARLSSON, B. *Beyond the C++ Standard Library: An Introduction to Boost*. Upper Saddle River: Addison-Wesley Professional, 2005. ISBN 0321133544.
- KRISHNASWAMY, R.; SIVARAJAN, K. Design of logical topologies: a linear formulation for wavelength-routed optical networks with no wavelength changers. *Networking, IEEE/ACM Transactions on*, v. 9, n. 2, p. 186–198, abr. 2001.
- LIMA, F. de O. *Um modelo eficiente para o projeto completo de redes ópticas*. Dissertação (Mestrado) — PPGEE - Programa de Pós Graduação em Engenharia Elétrica da Universidade Federal do Espírito Santo, Vitória, ES, 2010.
- LIMA, F. O. et al. Um Modelo Eficiente para o Projeto Completo de Redes Ópticas. In: *XLI Simpósio Brasileiro de Pesquisa Operacional*. Porto Seguro, BA: XLI SBPO, 2009.
- LIU, H.; TOBAGI, F. A. Physical topology design for all-optical networks. *Optical Switching and Networks*, Elsevier, 2008.

MICHALEWICZS; ZBIGNIEW. *Genetic Algorithms + Data Structures = Evolution Programs*. [S.l.]: Spring-Verlag, 1994.

MUKHERJEE, B. *Optical WDM networks*. Davis, CA - USA: Birkhäuser, 2006. (Optical networks series).

MUKHERJEE, B. et al. Some principles for designing a wide-area wdm optical network. *Networking, IEEE/ACM Transactions on*, v. 4, n. 5, p. 684–696, out. 1996.

MURTHY, C. S. R.; GURUSAMY, M. *WDM Optical Networks: Concepts, Design, and Algorithms*. Upper Saddle River: Prentice Hall, 2002.

NETO, P. L. de O. C. *Estatística*. 2. ed. São Paulo: Edgard Blücher, 2002.

PENG, I.-H.; CHEN, Y.-W.; WU, P.-W. A connectivity-based approach for the construction of hierarchical WDM network topology. *Springer - Photonic Network Communications*, 2008.

RAMAMURTHY, B. et al. Transparent vs. opaque vs. translucent wavelength routed optical networks. In: *Optical Fiber Communication Conference*. San Diego, CA , USA: OFC/IOOC, 1999. v. 1, p. 59–61.

RAMASWAMI, R.; SIVARAJAN, K. Design of logical topologies for wavelength-routed optical networks. *Selected Areas in Communications, IEEE Journal on*, v. 14, n. 5, p. 840–851, jun. 1996.

RAMASWAMI, R.; SIVARAJAN, K. N.; SASAKI, G. H. *Optical Networks: a practical perspective*. 3. ed. San Francisco: Morgan Kaufmann, 2009. (The Morgan Kaufmann series in networking).

SEGATTO, M. E. V. et al. Hybrid approaches for the design of mesh and hierarchical ring optical networks. In: LIMBERGER, H. G.; MATTHEWSON, M. J. (Ed.). SPIE, 2006. v. 6193, n. 1, p. 61931A. Disponível em: <<http://link.aip.org/link/?PSI/6193/61931A/1>>.

SHI, J.; FONSEKA, J. P. Hierarchical Self-Healing Rings. *IEEE/ACM Transactions on Networking*, Vol. 3, No. 6, 1995.

SILVA, M. et al. On Solving HSHR Networks. *6th Conference on Telecommunications - Conftele 2007*, 2007.

STERN, T. E.; ELLINAS, G.; BALA, K. *Multiwavelength Optical Networks: Architectures, Design, and Control*. New York, NY, USA: Cambridge University Press, 2008.

STUTZLE; DORIGO. ACO Algorithms for the Traveling Salesman Problem. *Université Libre de Bruxelles, Belgium*, 1999.

THOMADSEN, T. Hierarchical Network Design. *PHD tesis in Informatics and Mathematical Modelling - Technical University of Denmark*, 2005.

TORNATORE, M.; MAIER, G.; PATTAVINA, A. WDM Network Design by ILP Models Based on Flow Aggregation. *Networking, IEEE/ACM Transactions on*, v. 15, n. 3, p. 709–720, jun. 2007.

WALL, M. GALib: A C++ Library of Genetic Alrorithm Components. *MIT* - <http://lancet.mit.edu/ga/>, 2000.

XIAO, G.; LEUNG, Y.-W.; HUNG, K.-W. Two-Stage Cut Saturation Algorithm for Designing All-Optical Networks. *IEEE Transactions on Communications*, VOL. 49, No. 6, 2001.

XIN, Y.; ROUSKAS, G. N.; PERROS, H. G. On the physical and logical topology design of large-scale optical networks. *Lightwave Technology, Journal of*, v. 21, n. 4, p. 904–915, abr. 2003.

ZANG, H.; JUE, J. P.; MUKHERJEE, B. A Review of Routing and Wavelength Assignment Approaches for Wavelength Routed Optical WDM Networks. *Optical Networks Magazine*, SPIE/Baltzer Science Publishers, v. 1, p. 47–60, jan. 2000.

APÊNDICE A – Algoritmos Genéticos

Os algoritmos genéticos (*GAs - Genetic Algorithms*) são mecanismos de pesquisa baseados no processo de seleção natural e na combinação genética. Sua ideia principal é fazer com que indivíduos com certas características especiais, isto é, aqueles com mais afinidade com a função de estudo, sobrevivam e se combinem. Os GAs utilizam processos de escolha e combinação randômicos, seguindo algumas regras probabilísticas. Isto se deve ao fato de poder rastrear com menos vínculos um conjunto maior de indivíduos evitando cair em convergências a pontos de máximos e mínimos locais (GOLDBERG, 2002; MICHALEWICZS; ZBIGNIEW, 1994).

O GA foi primeiro proposto por Holland (1975). É um método de otimização que opera sobre uma população de indivíduos, que representam possíveis soluções do problema. A cada um é associada uma aptidão (*fitness*) para a solução do problema em questão. As populações de indivíduos são criadas e submetidas a operadores genéticos: de seleção, cruzamento (*crossover*) e mutação. Com base no *fitness* de cada indivíduo, estes operadores geram um processo de evolução natural dos mesmos. As possíveis respostas geradas para o problema são vistas como indivíduos que competirão entre si pela oportunidade de se reproduzirem. Os mais aptos têm maior chance de perpetuar parte de suas características, aumentando assim a probabilidade de se obter uma maior adaptação da população em geral (GOLDBERG, 2002). Algumas vantagens dos GA's são seu alto grau de adaptabilidade e paralelismo.

Como exemplo, considere um problema do caixeiro viajante ou TSP, um dos problemas de roteamento mais tradicionais e conhecidos de programação matemática (STUTZLE; DORIGO, 1999). O TSP é descrito por um conjunto de n cidades e uma matriz de distância entre elas, tendo o seguinte objetivo: o caixeiro viajante deve sair, de uma cidade chamada origem, e visitar cada uma das n cidades apenas uma única vez, retornando à cidade origem, percorrendo a menor distância possível, ou seja, deve ser encontrada uma rota fechada (ciclo hamiltoniano (CORMEN et al., 2002)) de comprimento mínimo que passe exatamente uma única vez por cada cidade. Na Figura A.1 temos um exemplo de uma solução para um TSP de quatro cidades. Uma solução para um TSP é uma combinação das n cidades que compõem o problema, e cada permutação desta é outra solução.

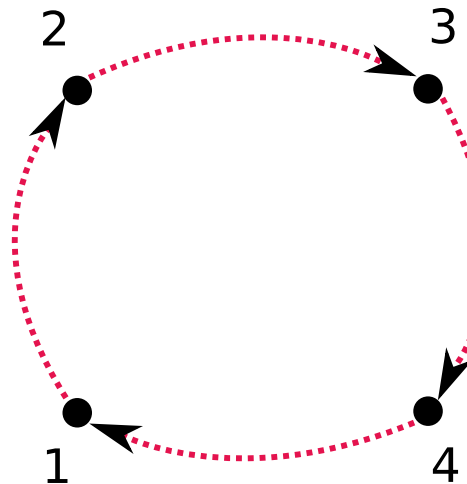


Figura A.1: Exemplo de uma solução para um TSP de quatro cidades.

Considere um algoritmo genético para um TSP de 6 cidades, com população inicial que contenha as sequencias [123654] e [213645]. A partir de um processo de seleção que leva em consideração o melhor desempenho de um indivíduo ao cumprimento da função objetivo (no caso minimizar o percurso entre as cidades), grupos de cromossomos par a par são escolhidos para reprodução. Num mecanismo de *crossover*, trocas de partes de seus genes são efetuadas e eventualmente ocorrem mutações possibilitando a variação numérica. Uma nova população com isso é gerada e o processo é reiniciado. Seleciona-se aleatoriamente um ponto de ruptura da sequencia (ponto de *crossing*), por exemplo, posição 3 : [123.654] e [213.645]. Permutando as duas partes dos cromossomos pais, cria-se uma próxima geração (*offspring*) de indivíduos : [123645] e [213654] que decodificados representam caminhos diferentes dos originais. Por sua vez se for efetuado uma mutação na sequencia [123645], permutando-se seus dois últimos alelos obtém-se [123654].

A codificação da informação em cromossomos é um ponto crucial durante o processo de modelagem do problema real, e junto com a função objetivo, liga o GA ao problema a ser resolvido. Se a codificação for feita de forma inteligente, essa já incluirá as características do problema e permitirá que se evitem testes de viabilidade de cada uma das soluções geradas. No exemplo, se utilizarmos a posição 4 como ponto de *crossing*, teríamos como resultado [123655] e [213644] que não representa uma solução para o TSP, e eles deverão ser descartados (lhes será atribuído um *fitness* que os exclua do processo evolutivo). Se a codificação não permitir esse tipo de anomalia, tais verificações serão desnecessárias, melhorando assim o desempenho do algoritmo.

O modelo de cromossomos mais utilizado é a representação binária, um vetor ou matriz de zeros e uns. Deste modo, uma codificação natural para uma topologia virtual seria tomar

a matriz de adjacências como cromossomo. Entretanto, o GA pode gerar qualquer matriz de zeros e uns de mesma ordem o que pode gerar topologias inviáveis. Na Tabela A.1 vemos uma matriz de adjacências e sua respectiva topologia virtual na Figura A.2.

0	1	0	0	0	0
0	0	1	0	0	0
0	0	0	1	0	0
0	0	0	0	1	0
0	0	0	0	0	1
1	0	0	0	0	0

Tabela A.1: Matriz de Adjacências de um anel

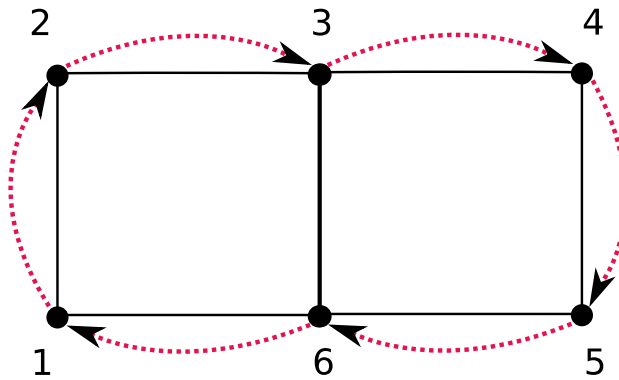


Figura A.2: Anel de 6 nós

Esta matriz poderia ser interpretada como um cromossomo binário 6x6, mas um indivíduo que tivesse entradas não nulas na diagonal não poderia representar uma topologia. Outro problema é a questão do grau lógico; as entradas não nulas do cromossomo, em cada linha e coluna, não podem exceder o grau lógico da rede. Além de outras restrições que podem ser consideradas, como o tipo de topologia (por exemplo, anéis). Mas, por outro lado, não importa ao GA como se codifica ou decodifica a informação dos parâmetros. Toda a informação relativa ao problema está contida na função objetivo do problema, que embute os módulos de codificação e decodificação dos parâmetros. Deste modo, os operadores genéticos devem ser implementados de forma a não gerarem soluções inviáveis, ou a função objetivo deverá estar apta a lidar com elas.

APÊNDICE B – Implementação de Algoritmos Genéticos

Diversas ferramentas para o auxílio na implementação de um GA encontram-se disponíveis atualmente. Destaca-se a biblioteca GALib ¹ (WALL, 2000), da linguagem de programação C++, que foi escolhida para trabalhar os GAs neste trabalho, por possuir vários algoritmos implementados, além de ter licença livre e código aberto. O que permite a utilização dos objetos prontos, ao mesmo tempo em que possibilita a livre adaptação e complementação.

A GALib é uma biblioteca multiplataforma (Unix, Linux, MacOS, Windows). Sua interface baseia-se em duas classes de objetos C++: o genoma (*genome*), e o GA (*genetic algorithm*). Cada objeto da classe *genome* é um indivíduo da população; uma solução para o problema. Já o objeto *genetic algorithm* define como a evolução será feita, utilizando uma função objetivo (definida pelo usuário) responsável pela decodificação do genoma e atribuição do *fitness* aos indivíduos (WALL, 2000).

Utilizando os operadores (implementados no *genome*) e as estratégias de seleção (implementados no *genetic algorithm*) são gerados novos indivíduos. Para um algoritmo genético simples, basta definir a representação (a maneira como uma solução será codificada em um genoma) a partir das classes de genoma existentes, definir a função objetivo e escolher um algoritmo genético, dentre os oferecidos pela biblioteca, e talvez passar mais alguns parâmetros, como critérios de parada e tamanho das populações. Novos genomas e operadores podem ser implementados pelo usuário para modificar um GA ou elaborar seu próprio algoritmo. Um GA contém as características gerais do algoritmo e dos operadores, a estratégia de evolução, os parâmetros de execução e o critério de parada, todos personalizáveis.

A GALib pode ser processada de maneira distribuída, evoluindo populações e/ou indivíduos paralelamente em múltiplos CPUs. Os parâmetros do algoritmo genético podem ser configurados em arquivo, em linha de comando, e/ou no próprio código. São suportados algoritmos com sobreposição de população (*Overlapping*), nos quais se pode definir o número de indivíduos (ou

¹<http://lancet.mit.edu/ga/>

uma porcentagem da população) a ser mantida a cada geração. A documentação da biblioteca inclui exemplos de outros algoritmos genéticos derivados, tais como um algoritmo genético com sub populações e outros que usam aglomeração determinística (*deterministic crowding*). Dentre os critérios internos de parada estão a convergência e o número de populações, que podem ser personalizados.

Os cromossomos podem ser construídos de qualquer tipo de dados de C++. Pode-se usar os tipos internos à biblioteca (*bit-string*, *array*, *list*, *tree*) ou criar um cromossomo baseado em seus próprios objetivos. Os tipos internos do cromossomo incluem *arrays* (arranjos), *list* (lista), *tree* (árvore), 1D, 2D, e 3D *arrays*, 1D, 2D, e 3D *binary string*. As listas e as árvores podem conter todo tipo de objeto em seus nós, bem como os arranjos em cada um de seus elementos. A biblioteca contém quatro tipos de genomas: *GAListGenome*, *GATreeGenome*, *GAArrayGenome*, e *GABinaryStringGenome*. Estas classes são derivadas da classe básica *GAGenome* e de uma classe de estrutura de dados como indicado por seus nomes (por exemplo, o *GAListGenome* é derivado das classes *GAList* e *GAGenome*). Por exemplo, se estiver tentando otimizar uma função que dependa de 5 números reais, use então como seu genoma um arranjo de uma dimensão de números reais com 5 elementos.

Há muitos tipos diferentes de algoritmos genéticos. A biblioteca *GAlib* inclui três tipos básicos: *simple*, *steady-state*, e *incremental*. Estes algoritmos diferem na maneira como criam indivíduos novos e substituem indivíduos velhos durante uma evolução. Esta biblioteca ainda fornece dois mecanismos preliminares que estendem as potencialidades dos objetos internos. Um recurso bastante utilizado é a definição de classes personalizadas com novas funções membro. Mas, se houver necessidade de fazer somente ajustes menores ao comportamento de uma classe *GAlib*, na maioria dos casos pode-se definir uma única função e dizer à classe existente para usá-la ao invés do *default*.

