

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

DANILO RICARDO BARBOSA DE ARAÚJO

PLANEJAMENTO DE REDES ÓPTICAS USANDO
INTELIGÊNCIA COMPUTACIONAL E CIÊNCIA DAS REDES



RECIFE, 2015.

DANILO RICARDO BARBOSA DE ARAÚJO

PLANEJAMENTO DE REDES ÓPTICAS USANDO
INTELIGÊNCIA COMPUTACIONAL E CIÊNCIA DAS REDES

Tese submetida ao Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal de Pernambuco como parte dos requisitos para obtenção do grau de Doutor em Engenharia Elétrica.

Área de Concentração: Comunicações.

ORIENTADOR: PROF. CARMELO JOSÉ ALBANEZ BASTOS FILHO, DOUTOR

CO-ORIENTADOR: PROF. JOAQUIM FERREIRA MARTINS FILHO, PhD

RECIFE, 2015.

©DANILO RICARDO BARBOSA DE ARAÚJO, 2015

Catálogo na fonte
Bibliotecária Margareth Malta, CRB-4 / 1198

A663p Araújo, Danilo Ricardo Barbosa de.
Planejamento de redes ópticas usando inteligência computacional e ciência das redes / Danilo Ricardo Barbosa de Araújo. - Recife: O Autor, 2015.
159 folhas, il., gráfs., tabs.

Orientador: Prof. Dr. Carmelo José Albanez Bastos Filho.
Coorientador: Prof. Dr. Joaquim Ferreira Martins Filho.
Tese (Doutorado) – Universidade Federal de Pernambuco. CTG.
Programa de Pós-Graduação em Engenharia Elétrica, 2015.
Inclui Referências e Apêndices.

1. Engenharia Elétrica. 2. Projeto de redes ópticas. 3. Inteligência computacional. 4. Ciência das redes. I. Bastos Filho, Carmelo José Albanez. (Orientador). II. Martins Filho, Joaquim Ferreira. (Coorientador). III. Título.

UFPE

621.3 CDD (22. ed.)

BCTG/2015-194



Universidade Federal de Pernambuco
Pós-Graduação em Engenharia Elétrica

PARECER DA COMISSÃO EXAMINADORA DE DEFESA DE
TESE DE DOUTORADO DE

DANILO RICARDO BARBOSA DE ARAÚJO

TÍTULO

**“PLANEJAMENTO DE REDES ÓPTICAS USANDO
INTELIGÊNCIA COMPUTACIONAL E CIÊNCIA DAS REDES”**

A comissão examinadora composta pelos professores: CARMELO JOSÉ ALBANEZ BASTOS FILHO, POLI/UPE; JOAQUIM FERREIRA MARTINS FILHO, DES/UFPE; RAUL CAMELO DE ANDRADE ALMEIDA JÚNIOR, DES/UFPE; MARCELO EDUARDO VIEIRA SEGATTO, DEE/UFES, MÔNICA DE LACERDA ROCHA, DEE/USP e FRANCISCO MADEIRO BERNARDINO JÚNIOR, POLI/UPE sob a presidência do primeiro, consideram o candidato **DANILO RICARDO BARBOSA DE ARAÚJO APROVADO.**

Recife, 21 de julho de 2015.

CECILIO JOSÉ LINS PIMENTEL
Coordenador do PPGEE

CARMELO JOSÉ ALBANEZ BASTOS FILHO
Orientador e Membro Titular Interno

MARCELO EDUARDO VIEIRA SEGATTO
Membro Titular Externo

JOAQUIM FERREIRA MARTINS FILHO
Coorientador e Membro Titular Interno

MÔNICA DE LACERDA ROCHA
Membro Titular Externo

**RAUL CAMELO DE ANDRADE ALMEIDA
JÚNIOR**
Membro Titular Interno

**FRANCISCO MADEIRO BERNARDINO
JÚNIOR**
Membro Titular Externo

À minha mãe pela vida, dedicação e educação, e à minha esposa pelo incentivo e companheirismo.

Agradecimentos

À minha mãe pelo exemplo de vida, pela dedicação integral aos filhos e por ter colocado a educação sempre em primeiro lugar. À minha esposa pelo companheirismo durante esta jornada.

Ao professor e orientador Carmelo José Albanez Bastos Filho pela oportunidade, orientação, conselhos, confiança e compreensão.

Ao professor e co-orientador Joaquim Ferreira Martins Filho pela orientação e acompanhamento durante todo o período.

Aos demais professores, aos mestrandos e aos doutorandos que fazem parte do grupo de pesquisa em redes ópticas da UFPE, pelas discussões de elevado nível que sempre ocorreram durante as reuniões.

Aos colaboradores que fazem parte da parceria UFPE-Padtec pela oportunidade de colocar em prática alguns aspectos teóricos sob o formato de uma ferramenta de planejamento de redes.

A todos os demais professores que contribuíram para a minha formação.

Resumo da Tese apresentada à UFPE como parte dos requisitos necessários para a obtenção do grau de Doutor em Engenharia Elétrica.

Planejamento de Redes Ópticas Usando Inteligência Computacional e Ciência das Redes

Danilo Ricardo Barbosa de Araújo

Julho/2015

Orientador: Prof. Carmelo José Albanez Bastos Filho, Doutor.

Co-orientador: Prof. Joaquim Ferreira Martins Filho, PhD.

Área de Concentração: Comunicações.

Palavras-chaves: Projeto de Redes Ópticas, Inteligência Computacional, Ciência das Redes.

Número de páginas: 159.

A infraestrutura de transporte para Internet e outros serviços de telecomunicações que exigem elevada taxa de transmissão é constituída basicamente por redes de comunicações ópticas. Em redes com estabelecimento dinâmico de requisições, algumas das requisições podem não ser atendidas devido à indisponibilidade de um caminho óptico, decorrente da ausência de um canal ou de restrições de Qualidade de Transmissão (QoT). A probabilidade de bloqueio de requisições (PB) pode ser usada para medir o desempenho deste tipo de rede. PB pode ser estimada por expressões analíticas ou por simuladores de eventos discretos. Expressões analíticas possuem baixo custo computacional, mas são aplicadas apenas em cenários muito simples. Simuladores facilitam a estimativa de PB em cenários mais complexos, mas apresentam elevado custo computacional, pois exigem um grande número de requisições para uma estimativa precisa. Algoritmos evolucionários foram usados em trabalhos anteriores sobre planejamento de redes, mas eles apresentam elevado custo computacional, pois ativam o estimador de PB diversas durante o processo. Nesta Tese é proposta uma nova abordagem para planejamento de redes baseada em conceitos de Ciência das Redes e de Inteligência Computacional. São propostas novas métricas, um estimador de PB baseado em redes neurais e um novo algoritmo de planejamento que usa o conceito de modelos substitutos. De acordo com os resultados obtidos, o novo método de estimativa de PB possui um tempo até 72 vezes menor do que o obtido por um simulador e o erro de estimativa é da ordem de 10^{-4} , para uma rede óptica de 18 nós. Diversos cenários de projeto de redes foram analisados e o novo algoritmo de planejamento oferece uma redução do tempo de execução de até 88% sobre as propostas anteriores, baseadas em algoritmos evolucionários e simuladores de rede.

Abstract of Thesis presented to UFPE as a partial fulfillment of the requirements for the degree of Doctor in Electrical Engineering

Design of Optical Networks by Using Computational Intelligence and Network Science

Danilo Ricardo Barbosa de Araújo

July/2015

Supervisor: Prof. Carmelo José Albanez Bastos Filho, Doctor.

Co-supervisor: Prof. Joaquim Ferreira Martins Filho, PhD.

Area of Concentration: Communications.

Keywords: Optical Networks Design, Computational Intelligence, Network Science.

Number of pages: 159.

The transport infrastructure for the Internet and other telecommunications services that require high transmission rate is comprised basically of optical communication networks. If networks with dynamic traffic are used, some requests could not be established due to the absence of an optical channel or due to restrictions in the Quality of Transmission (QoT). The blocking probability (BP) can be used to measure the performance of this kind of networks. BP can be estimated by analytical expressions or by discrete event simulators. Analytical expressions provide low computational cost, but they can be used only in very simple scenarios. Simulators facilitate the estimation of BP in more complex scenarios, but they present a high computational cost due to the large number of calls which is necessary to provide an accurate estimation. Evolutionary algorithms were used in previous works about design of optical networks, but they present high computational cost, since they activate the BP estimator several times to plan a single network. This Thesis proposes a new approach to design optical networks based on Network Science and Computational Intelligence. We propose new metrics, a new method to estimate BP based on neural networks and a new planning approach based on surrogate models. From the results, the proposed BP estimator presents a speedup of 72 over a network simulator and presents estimation errors in the order of 10^{-4} , for an optical network with 18 nodes. Different scenarios for optical network design were evaluated and the new proposal offers a reduction in processing time up to 88% when it is compared with previous approaches, that are based on evolutionary algorithms and network simulators.

Lista de Abreviaturas e Siglas

AAH	<i>Asynchronous Amplitude Histogram</i> - Histograma de amostragem assíncrona
ACP	Análise de Componentes Principais - Técnica para análise da relação entre variáveis
AE	Algoritmo Evolucionário - Família de meta-heurísticas
AEMO	Algoritmo Evolucionário Multi-objetivos - Sub-classificação de AE
APL	<i>Average Path Length</i> - Comprimento médio dos caminhos mínimos
ASE	<i>Amplified Spontaneous Emission</i> - Emissão espontânea amplificada
BA	Barabási-Albert - Modelo de geração de grafos <i>scale-free</i>
BER	<i>Bit Error Rate</i> - Taxa de erro de bit
BP	<i>Blocking Probability</i> - Probabilidade de Bloqueio
CAC	Controle de Admissão de Chamadas - Módulo para admissão de um pedido de conexão
CAPEX	<i>Capital Expenditure</i> - Custo para implantação de uma rede
CC	<i>Clustering Coefficient</i> - Coeficiente de agrupamento de um grafo
CD	<i>Crowding Distance</i> - Métrica de diversidade baseada na distância entre as soluções
CMS	Cascata de modelos substitutos - Mecanismo para avaliação de soluções
CR	Concentração de rotas - Métrica para avaliação de rede
CS	<i>Coverage of two Set</i> - Métrica para avaliar a convergência relativa entre duas frentes de Pareto
DFT	<i>Discrete Fourier Transform</i> - Transformada Discreta de Fourier
DLE	<i>Dynamic Lightpath Establishment</i> - Estabelecimento dinâmico de caminhos ópticos
DWDM	<i>Dense Wavelength Division Multiplexing</i> - Multiplexação por divisão densa em comprimento de onda
EDFA	<i>Erbium Doped Fiber amplifier</i> - Amplificador à fibra dopada com Érbio
ER	Erdos-Renyi - Modelo de geração de grafos aleatórios
F	<i>Noise Factor</i> - Fator de Ruído
FWM	<i>Four Wave Mixing</i> - Mistura de quatro ondas
FZC	<i>First Zero Component</i> - Primeiro componente a assumir valor zero
HV	<i>Hypervolume</i> - Indicador de qualidade para AEMO
HVC	<i>Highest Value Component</i> - Componente que apresenta o maior valor
IF	<i>Isolation Factor</i> - Fator de isolamento nas portas de um ROADM
IMM	<i>Interactive Matrix Methodology</i> - Metodologia iterativa com matrizes

LM	Levenberg-Marquardt - Método de treinamento de RNA
MO-	<i>Multi-objective Minimum Spanning Tree</i> - Versão multiobjetivos do problema
MST	MST
MS	<i>Maximum Spread</i> - Métrica para avaliar o espalhamento máximo em frentes de Pareto
MSE	<i>Mean Squared Error</i> - Erro médio quadrático
MST	<i>Minimum Spanning Tree</i> - Árvore geradora mínima
NC	<i>Natural Conectivity</i> - Conectividade natural
NF	<i>Noise Figure</i> - Figura de ruído
NFSNET	<i>National Science Foundation Network</i> - Rede de transporte Americana
NSGA	<i>Non-dominated Sorting Genetic Algorithm</i> - Algoritmo de otimização multiobjetivos
OADM	<i>Optical Add and Drop Multiplexer</i> - Multiplexador óptico de adição e remoção de comprimentos de onda
OEO	Óptico-Eletrônico-Óptico - Conversão do sinal óptico em elétrico e depois em óptico
OSNR	<i>Optical Signal to Noise Ratio</i> - Relação sinal-ruído óptica
OTN	<i>Optical Transport Network</i> - Rede de transporte óptica
PB	Probabilidade de bloqueio - Indicador de qualidade de redes sujeitas a DLE
PC	<i>Principal Component</i> - Componente principal, em ACP
PMD	<i>Polarization Mode Dispersion</i> - Dispersão por modo de polarização
PTD	<i>Physical Topology Design</i> - Projeto de topologia física
QoT	<i>Quality of Transmission</i> - Qualidade de Transmissão
RC	<i>Request of Calls</i> - Pedidos de conexões a uma rede
RD	<i>Residual Dispersion</i> - Dispersão cromática residual
RLM	Regressão Linear múltipla - Método de regressão
RNA	Rede Neural Artificial - Família de técnicas de aprendizagem de máquina
RND	<i>Random</i> - Aleatório
ROADM	<i>Reconfigurable optical add-drop multiplexer</i> - Tipo de OADM que permite configurar remotamente a comutação do tráfego em redes WDM
RSS	<i>Residual Sum of Squares</i> - Soma dos quadrados dos resíduos
RWA	<i>Routing and Wavelength Assignment</i> - Roteamento e atribuição de comprimentos de onda
SDH	<i>Synchronous Digital Hierarchy</i> - Esquema de multiplexação TDM
SF	<i>Scale-free</i> - Livre de escala
SLE	<i>Static Lightpath Establishment</i> - Estabelecimento estático de caminhos ópticos
SONET	<i>Synchronous Optical Network</i> - Rede óptica síncrona
SP	<i>Shortest Path</i> - Caminho mais curto
SW	<i>Small-world</i> - Efeito de mundo pequeno
VTD	<i>Virtual Topology Design</i> - Projeto de topologia lógica
WDM	<i>Wavelength Division Multiplexing</i> - Multiplexação por divisão em comprimento de onda
WRON	<i>Wavelength Routed Optical Network</i> - Rede óptica roteada por comprimento de onda
WS	Watts-Strogatz - Método de geração de grafos com efeito SW

WSS *Wavelength selective switches* - Comutadores de comprimento de onda seletivos

WS-T Watts-Strogatz guiado por Tráfego - Método de geração de grafos que considera uma matriz de tráfego

Lista de Símbolos

$OSNR_{in}$	Relação sinal-ruído óptica na entrada de um dispositivo
$OSNR_{out}$	Relação sinal-ruído óptica na saída de um dispositivo
$OSNR_{TH}$	Relação sinal-ruído óptica de limiar para fins de QoT
P_S	Potência do sinal
P_N	Potência do ruído
F	Fator de ruído
NF	Figura de ruído
$\Delta\tau$	Alargamento temporal do pulso óptico
PB	Probabilidade de bloqueio
G	O grafo de uma rede
\mathcal{N}	Conjunto de nós de uma rede
\mathcal{E}	Conjunto de enlaces de uma rede
n	Número de nós de uma rede
e	Número de enlaces de uma rede
e_{max}	Número máximo de enlaces em uma rede
\mathcal{A}	Matriz de adjacências de uma rede
k_l	Conectividade por enlace
q_{km}	Densidade de implantação de fibras
q	Densidade de enlaces de uma rede
d	Grau de um nó
\hat{d}	Grau médio de nós em uma rede
r	Coefficiente de assortatividade
$SP_{i,j}$	Caminho mais curto entre os nós i e j
$diam(G)$	Diâmetro de uma rede
\bar{c}	Comprimento médio dos caminhos mínimos
NC	Conectividade natural de um grafo
cc_i	Coefficiente de agrupamento do nó i
CC	Coefficiente de agrupamento de uma rede
$I(G)$	Entropia dos graus dos nós de uma rede
cc_i	Coefficiente de agrupamento do nó i
\mathcal{D}	Matriz de grau de uma rede
\mathcal{L}	Matriz Laplaciana de uma rede
rp	Probabilidade de religação de enlaces para o modelo WS
ρ	Raio espectral
λ_i	i -ésimo autovalor da matriz Laplaciana de uma rede

λ_{n-1}	Conectividade algébrica
$f_\lambda(t)$	Função de densidade dos autovalores
rp	Probabilidade de religação para o modelo de redes WS
$I(\hat{\mathcal{F}})$	Entropia dos coeficientes da DFT dos autovalores do Laplaciano de uma rede
τ	Expoente usado na conexão preferencial do modelo BA
δ	Fator de isolamento entre as portas de um ROADM
W	Número máximo de comprimentos de onda em uma rede óptica, por enlace
L	Carga de uma rede, em erlangs
p	Quantidade de variáveis de entrada consideradas em um modelo de regressão
M	Número de neurônios da camada escondida de uma MLP
B_0	Largura de banda do filtro
Δf	Espaçamento entre canais
α	Coefficiente de atenuação da fibra
L_{Mux}	Perda no multiplexador
L_{Demux}	Perda no demultiplexador
L_{Switch}	Perda no switch óptico
Ω	Espaço de busca em um problema de otimização
\mathcal{P}^*	Conjunto Pareto ótimo
\vec{X}	Vetor de variáveis de decisão para o problema de otimização
P	Número de indivíduos na população de um algoritmo evolucionário
P_M	Probabilidade de mutação
P_C	Probabilidade de cruzamento
I_{max}	Número máximo de iterações em um processo iterativo
σ_{PL}	Desvio padrão dos caminhos mínimos
$\overline{\Delta OSNR}$	Margem média de OSNR
$\sigma_{\Delta OSNR}$	Desvio padrão de $\Delta OSNR$
T	Matriz de tráfego de uma rede

Lista de Figuras

1.1	Exemplos de diferentes configurações de rede para uma mesma demanda de projeto físico.	23
3.1	Exemplo de cálculo da métrica CC para uma rede de 10 nós.	39
3.2	Exemplos de redes criadas por quatro procedimentos de geração diferentes. . .	42
3.3	Valores dos componentes pares e ímpares da DFT dos autovalores da matriz Laplaciana de uma rede livre de escala.	44
3.4	Comparação de FZC (a) e HVC (b) de redes com 100 nós.	47
3.5	Comparação de FZC (a) e HVC (b) de redes com 200 nós.	47
3.6	Comparação de FZC (a) e HVC (b) de redes com 400 nós.	47
3.7	Gráfico de <i>box-plot</i> para análise da variação da métrica FZC para um conjunto de 100 amostras de cada tipo e para $q = 0, 10$	48
3.8	Autovalores da matriz Laplaciana e componentes pares da DFT dos autovalores para redes com mesmo número de nós de densidade da rede “Sociopatterns-Hypertext 2009”.	50
3.9	Autovalores da matriz Laplaciana e componentes pares da DFT dos autovalores para redes com mesmo número de nós de densidade da rede “Caenorhabditis elegans metabolic”.	51
3.10	Métrica $I(\hat{\mathcal{F}})$ versus densidade para redes k -regulares, redes <i>scale-free</i> geradas pelo modelo BA, redes aleatórias geradas pelo modelo ER e redes <i>small-world</i> geradas pelo modelo WS.	55
3.11	Entropia do grau dos nós e entropia dos autovalores de \mathcal{L} versus densidade para redes <i>scale-free</i> , redes aleatórias e redes <i>small-world</i> , de 100 nós.	56
3.12	Box-plot para 30 redes diferentes com $n = 100$ e $d = 0,04$ (a); $d = 0,30$ (b) e $d = 0,80$ (c), respectivamente.	57
3.13	A rede NSFNet original, com 14 nós.	62
3.14	Comparação entre duas topologias semelhantes, mas que apresentam diferentes valores de CR.	63

3.15	Gráficos de dispersão para avaliação da correlação entre a métrica proposta CR e a métrica de centralização de Freeman [1].	64
4.1	Estrutura de um enlace em uma rede WDM considerada pelo SIMTON.	69
4.2	Arquitetura de um nó transparente para as simulações do SIMTON.	70
4.3	Fluxograma de execução do CAC transparente.	72
4.4	Fluxograma com os passos da metodologia proposta nesta Tese para obtenção de um estimador de PB baseado em RNA.	73
4.5	RNA com uma camada escondida para obtenção de PB.	74
4.6	Uma visão geral do analisador de redes proposto.	76
4.7	Topologias de fibra de três diferentes redes de transporte.	82
4.8	\overline{MSE} de PB para os “melhores modelos” usando diferentes números de (p) entradas e dados do SIMTON A e B, para uma rede sujeita a um tráfego de 200 erlangs.	88
4.9	Comparação de duas redes WDM similares e o impacto das métricas na avaliação do desempenho.	89
4.10	PB em função da carga da rede para o SIMTON e para a RNA treinada com 100 e 200 erlangs.	89
4.11	Estimativa de PB usando três diferentes abordagens e considerando variação da carga da rede, para o problema Tipo A.	93
4.12	\overline{MSE} e tempo de treinamento para diferentes números de redes usadas no treinamento da RNA para o problema Tipo A.	94
4.13	\overline{MSE} e tempo de treinamento para diferentes números de redes usadas no treinamento da RNA para o problema Tipo B.	94
4.14	<i>Box-plots</i> da média do erro médio quadrático, considerando 20 avaliações para cada método de estimativa de PB, para redes ARNES, FLTG e MEDIANET.	97
5.1	Um problema com duas funções objetivo: custo e atraso da rede. A frente de Pareto ou superfície de compromisso está destacada pela linha que une os pontos $A - E$	101
5.2	Ilustração do mecanismo de seleção utilizado pelo NSGAIL.	105
5.3	Modelo para seleção de sobreviventes usado no NSGAIL.	105
5.4	Exemplo de compromisso entre fidelidade das estimativas e eficiência de recursos para modelos substitutos em avaliação de redes ópticas.	107
5.5	Comparação de CS para diferentes algoritmos de inicialização, para o cenário de tráfego uniforme e para o caso NSFNet.	113
5.6	Comparação de CS para diferentes algoritmos de inicialização, para o cenário de tráfego não-uniforme e para o caso NSFNet.	114

5.7	Frentes de Pareto obtidas para diferentes algoritmos de inicialização de projetos de redes ópticas para o caso NSFNet.	114
5.8	Frentes de Pareto obtidas para diferentes algoritmos de inicialização de projetos de redes ópticas para MEDIANET, ARNES e FLTG.	116
5.9	Ilustração da metodologia proposta para tratamento da cascata de modelos substitutos.	120
5.10	Indicadores de qualidade aplicados às frentes de Pareto obtidas após 15 minutos de execução para a nova proposta evolucionária e para o EA-NS. O cenário considerado foi o de tráfego uniforme para a NSFNet.	125
5.11	Frentes de Pareto obtidas após 15 minutos de execução do EA-NS e da nova proposta e uma frente de Pareto de referência. O cenário considerado é o de tráfego uniforme para a NSFNet.	126
5.12	Indicadores de qualidade aplicados às frentes de Pareto obtidas após 15 minutos de execução para a nova proposta evolucionária e para o EA-NS. O cenário considerado foi o de tráfego não-uniforme para a NSFNet.	126
5.13	Frentes de Pareto obtidas após 15 minutos de execução do EA-NS e da nova proposta e uma frente de Pareto de referência. O cenário considerado é o de tráfego não-uniforme para a NSFNet.	127
5.14	Frentes de Pareto obtidas após 15 minutos de execução do EA-NS e da nova proposta para redes ARNES, FLTG e MEDIANET.	130

Lista de Tabelas

3.1	Comparação de propriedades topológicas de alguns modelos de rede para grafos com 50 nós, 100 arestas e $rp = 5\%$	43
3.2	Comparação entre redes reais e redes geradas por modelos canônicos que possuem o mesmo número de nós e densidade das redes originais.	49
3.3	Redes reais e redes WS, ER e BA equivalentes.	58
3.4	Matriz C para a rede NSFNet original.	62
4.1	Variáveis independentes consideradas para estimativa de PB.	78
4.2	Parâmetros de simulação das redes WDM.	79
4.3	Classes de intervalos de PB para criar a base de dados de redes ópticas.	79
4.4	Variáveis independentes incluídas na etapa final do estudo sobre a estimativa de PB.	84
4.5	Classes de intervalos de PB para criar a base de dados de redes ópticas na etapa final do estudo.	85
4.6	Matriz de correlação das variáveis independentes (vide Tabela 4.1).	85
4.7	Percentual da variação total nos dados explicada pelas componentes principais da matriz de correlação de variáveis que influenciam a PB.	86
4.8	Correlação entre componentes principais e variáveis originais para o problema de regressão de PB.	87
4.9	Melhores modelos de regressão com p variáveis para estimativa de PB para a rede NSFNet.	87
4.10	Melhores modelos de regressão para estimativa de PB para redes FLTG, ARNES e MEDIANET.	90
4.11	Comparação entre duas diferentes arquiteturas de RNA para estimativa de PB dos estudos de FLTG, ARNES e MEDIANET.	92
4.12	Comparação entre o tempo médio de avaliação para simulações de rede e o tempo médio de avaliação para RNAs.	95
5.1	Variáveis de decisão para planejamento das redes ópticas consideradas.	111

5.2	Parâmetros de simulação para o projeto das redes considerados no arranjo experimental.	111
5.3	Média do indicador <i>coverage set</i> , para tráfego uniforme e para os cenários FLTG, ARNES e MEDIANET.	115
5.4	Média do indicador <i>coverage set</i> , para tráfego não-uniforme e para os cenários FLTG, ARNES e MEDIANET.	115
5.5	Modelos Substitutos para estimativa de PB de redes ópticas.	119
5.6	Parâmetros do algoritmo evolucionário usado para projeto de redes.	123
5.7	Informações detalhadas de algumas redes que estão destacadas na Figura 5.13.	128
5.8	Um resumo do uso de cada modelo substituto em um fluxo completo de execução da nova proposta evolucionária baseada em CMS para a NSFNet.	128
5.9	Média e desvio padrão dos indicadores de qualidade, para tráfego uniforme e não-uniforme e para redes FLTG, ARNES e MEDIANET.	130
6.1	Resumo das publicações e contribuições da Tese.	135
A.1	Melhores modelos de regressão com p variáveis para estimativa de PB para o problema Tipo A e para a rede MEDIANET.	150
A.2	Melhores modelos de regressão com p variáveis para estimativa de PB para o problema Tipo B e para a rede MEDIANET.	151
A.3	Melhores modelos de regressão com p variáveis para estimativa de PB para o problema Tipo A e para a rede ARNES.	151
A.4	Melhores modelos de regressão com p variáveis para estimativa de PB para o problema Tipo B e para a rede ARNES.	151
A.5	Melhores modelos de regressão com p variáveis para estimativa de PB para o problema Tipo A e para a rede FLTG.	152
A.6	Melhores modelos de regressão com p variáveis para estimativa de PB para o problema Tipo B e para a rede FLTG.	152
B.1	Matriz de tráfego usada para o cenário de tráfego não-uniforme para a rede NSFNet.	154
B.2	Matriz de tráfego usada para o cenário de tráfego não-uniforme para a rede MEDIANET.	155
B.3	Matriz de tráfego usada para o cenário de tráfego não-uniforme para a rede ARNES (continua).	156
B.4	Matriz de tráfego usada para o cenário de tráfego não-uniforme para a rede ARNES (conclusão).	157
B.5	Matriz de tráfego usada para o cenário de tráfego não-uniforme para a rede FTLG (continua).	158

B.6	Matriz de tráfego usada para o cenário de tráfego não-uniforme para a rede	
	FTLG (conclusão).	159

Sumário

1	Introdução	21
1.1	Organização do Documento	24
2	Revisão da Literatura em Análise e Projeto de Redes Ópticas	26
2.1	Análise de Redes Ópticas	27
2.1.1	Modelo para Avaliação do Custo de uma Rede Óptica	27
2.1.2	Análise da Qualidade do Sinal em Enlaces Ópticos	29
2.1.3	Análise de Desempenho em Redes Ópticas	31
2.2	Planejamento de Redes Ópticas	33
3	Ciência das Redes	36
3.1	Características Topológicas de Redes	37
3.2	Modelos Canônicos de Redes	41
3.3	Proposta das Métricas FZC e HVC	43
3.3.1	Descrição	43
3.3.2	Arranjo Experimental	45
3.3.3	Resultados	46
3.3.4	Considerações Finais sobre FZC e HVC	52
3.4	Proposta da Métrica $I(\hat{\mathcal{F}})$	52
3.4.1	Descrição	52
3.4.2	Arranjo Experimental	54
3.4.3	Resultados	55
3.4.4	Considerações Finais sobre $I(\hat{\mathcal{F}})$	57
3.5	Proposta de uma Métrica baseada em Concentração de Rotas (CR)	58
3.5.1	Descrição	58
3.5.2	Arranjo Experimental	61
3.5.3	Resultados	62
3.5.4	Considerações Finais sobre CR	65

4	Análise de Redes Ópticas	67
4.1	Simulador de Eventos Discretos	68
4.1.1	Modelo de Camada Física	69
4.1.2	Geração de Chamadas	70
4.1.3	Controle de Admissão de Chamadas (CAC) e Estimativa de PB	71
4.2	Proposta de uma Nova Metodologia para Análise de Redes Ópticas	72
4.2.1	Descrição	73
4.2.2	Arranjo Experimental	76
4.2.3	Resultados Obtidos	84
4.2.4	Considerações Finais sobre Estimativa de PB Usando RNA	97
5	Planejamento de Redes Ópticas	99
5.1	Otimização Multiobjetiva	100
5.1.1	Conceitos Básicos	101
5.1.2	Métricas para Avaliação de Desempenho de AEMO	102
5.1.3	O Algoritmo NSGAI	103
5.1.4	Otimização Assistida por Modelos Substitutos	106
5.2	O Problema de Planejamento de Redes Ópticas	107
5.3	Um Novo Método para Inicialização de Projetos de Redes Ópticas	109
5.3.1	Descrição	109
5.3.2	Arranjo Experimental	110
5.3.3	Resultados	113
5.4	Cascata de Modelos Substitutos para Avaliação de Redes em Algoritmos de Otimização	117
5.4.1	Descrição	117
5.4.2	Arranjo Experimental	122
5.4.3	Resultados Obtidos	124
5.5	Considerações Finais sobre a Nova Abordagem para Planejamento de Redes	131
6	Conclusões	133
6.1	Resumo Quantitativo das Contribuições da Tese	136
6.2	Sugestões para Trabalhos Futuros	136
	Referências	138
	Apêndice A Combinações de Variáveis Independentes para Estimativa de PB150	
	Apêndice B Matriz de Tráfego das Redes Analisadas	153

Capítulo 1

Introdução

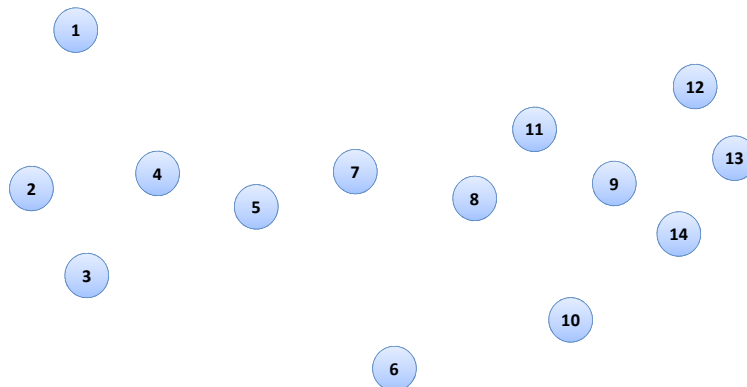
O crescente aumento de demanda por taxa de transmissão, relacionado em sua maior parte pelo crescimento contínuo de tráfego de Internet, tem sido suprido nos últimos anos primariamente por comunicações por fibra óptica [2]. Estudos recentes demonstram que o crescimento do tráfego de Internet continua em expansão e deve ser aumentado em um fator de 1.000 nos próximos 20 anos [3]. Com o objetivo de continuar suprindo demandas crescentes por taxa de transmissão, diversas ações são necessárias em conjunto, tais como: uso sistemático de redes dinâmicas, melhorias em eficiência espectral, expansão da banda de transmissão, dentre outras [3].

As redes ópticas atualmente em operação são muitas vezes semi-estáticas e a configuração de novas conexões geralmente requer uma interferência manual de um operador. Estas conexões são mantidas por meses ou anos. Estudos demonstram que o uso de redes ópticas dinâmicas pode melhorar o desempenho global do sistema, pois as conexões são mantidas apenas durante o tempo mínimo necessário e em seguida os recursos da rede são liberados para alocação de novas conexões [4, 3]. Considerando as complexidades adicionais envolvidas em redes ópticas dinâmicas quando comparadas com redes estáticas equivalentes, justifica-se o desenvolvimento de ferramentas próprias para análise e projeto de redes que considerem as peculiaridades deste cenário.

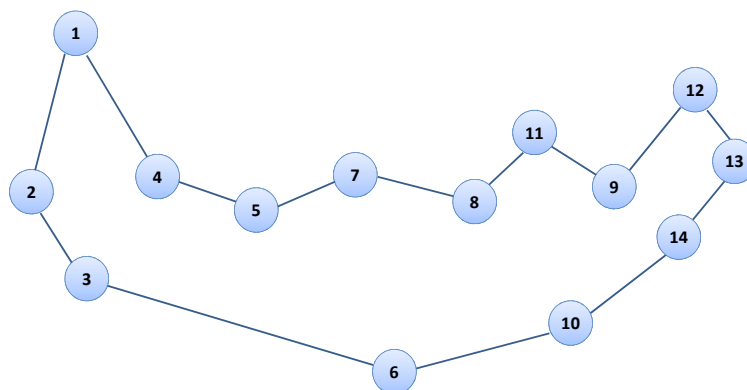
Diversas soluções já foram propostas para análise de enlaces ópticos [5, 6] e para análise de redes ópticas [7, 8, 9, 10, 11, 12]. Entretanto, diferente das ferramentas de análise, pouco tem sido feito no sentido do desenvolvimento de ferramentas próprias para auxiliar o projeto de redes ópticas. O projeto de uma rede óptica pode ser classificado em duas principais tarefas: projeto da topologia física e dos equipamentos usados na rede (PTD - *Physical Topology Design*); e projeto da topologia lógica (VTD - *Virtual Topology Design*) [2]. VTD consiste na maior parte das vezes em embarcar os caminhos ópticos sobre uma topologia física existente. O projeto físico é um problema computacionalmente custoso, sendo classificado como um problema NP-difícil e portanto não é possível obter uma solução ótima em tempo aceitável

para redes grandes [13]. Diversos trabalhos consideram a busca por soluções subótimas usando heurísticas e meta-heurísticas. A maior crítica aos primeiros estudos sobre PTD de redes ópticas é o fato de focarem apenas em redes estáticas ou usarem uma modelagem do problema que otimiza apenas um dos objetivos da rede, como custo ou uso de comprimentos de onda [14, 15, 16, 17, 18, 19, 20]. Trabalhos recentes demonstraram que é possível realizar PTD para redes ópticas considerando uma modelagem baseada em objetivos conflitantes e um cenário de tráfego dinâmico que considera efeitos da camada física para estimativa da probabilidade de bloqueio das redes avaliadas pelo algoritmo de planejamento [21, 22]. Contudo, apesar destes trabalhos usarem meta-heurísticas, o tempo total necessário para determinação de um projeto adequado de rede óptica pode chegar a vários dias. Este tempo está relacionado principalmente com as chamadas frequentes a um simulador de redes que considera os efeitos da camada física para testar diversos projetos que são criados iterativamente [21, 22].

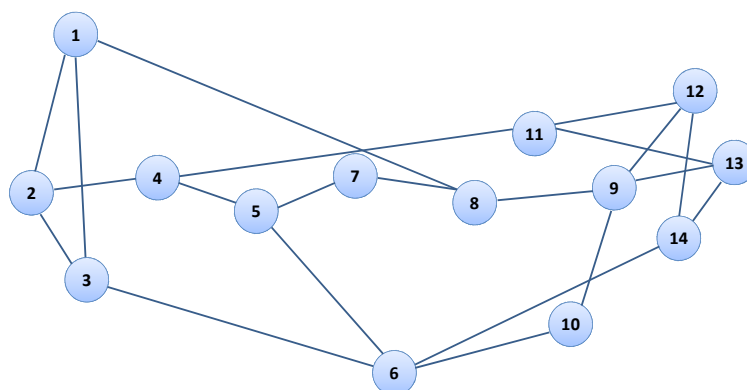
Para um melhor entendimento sobre o problema de projeto físico de redes ópticas semelhante ao que é investigado em [21, 22], considere a necessidade de construção de um exemplo de rede WDM. Esta rede será criada a partir da localização geográfica dos nós ilustrados na Figura 1.1(a), de um algoritmo de RWA e de uma demanda de tráfego. Com base nestes dados de entrada, podem ser construídas diversas topologias que apresentam compromissos distintos entre o desempenho da rede e o custo de implantação (CAPEX - *Capital Expenditure*). Considere que o desempenho será avaliado por meio de uma estimativa da PB da rede e o custo será calculado por um modelo de custo que fornece valores em unidades monetárias genéricas (u.m.) [21]. Suponha que o projeto esteja relacionado com uma demanda de tráfego uniforme de 100 erlangs e seja usado um algoritmo de menor caminho em conjunto com o *first fit*. Também são usados amplificadores ópticos EDFAs com os ganhos ajustados para compensar as perdas nas fibras ópticas e nos equipamentos dos nós da rede. A Figura 1.1(b) ilustra um exemplo de topologia de rede em anel. Caso seja adotado um número máximo de 40 comprimentos de onda por enlace e um modelo de ROADM que apresente um fator de isolamento de 40 dB na rede da Figura 1.1(b), esta rede apresenta $PB = 0,27$ quando analisada por um simulador de rede que considera os efeitos da camada física [11]. A rede da Figura 1.1(c) representa uma possível configuração de rede em malha e contém diversos enlaces adicionais quando comparada à rede em anel. Caso a topologia da Figura 1.1(c) seja submetida às mesmas condições apresentadas para a rede da Figura 1.1(b), a PB encontrada para esta nova topologia será $PB = 2,6 \cdot 10^{-4}$. Se estas redes forem analisadas usando o modelo de custo proposto por [21], será obtido para a rede da Figura 1.1(b) $CAPEX = 2.888$ u.m. e para a rede da Figura 1.1(c) $CAPEX = 4.747$ u.m. Caso o número de comprimentos de onda máximo por enlace seja alterado no projeto da rede e caso o modelo de ROADM usado seja modificado, o desempenho e o custo das redes também serão afetados. Para se obter um conjunto de soluções que apresentem uma relação ótima de compromisso entre custo e desempenho da rede seria necessária a análise de um grande número de redes WDM relacionadas



(a) Disposição dos nós da rede para um projeto físico de rede WDM.



(b) Proposta de topologia de rede em anel.



(c) Proposta de topologia de rede em malha.

Figura 1.1: Exemplos de diferentes configurações de rede para uma mesma demanda de projeto físico.

com todas as possibilidades de variação da topologia física e demais especificações do projeto. Considerando uma rede de 14 nós e a necessidade de informar o número de comprimentos de onda (entre 1 e 40) e um modelo de ROADM (considerando 5 modelos diferentes), a quantidade total de possibilidades é de $2^{91} \cdot 5 \cdot 40 \approx 5 \cdot 10^{29}$. Considerando que o simulador de redes usado neste exemplo apresenta um tempo médio para simulação de cada rede superior a 1 segundo [21], seriam necessários pelo menos $1,59 \cdot 10^{22}$ anos para avaliar todas as possibilidades. Portanto, o problema de projeto físico de redes WDM precisa ser solucionado por técnicas de otimização que apresentem soluções com bons compromissos entre desempenho e custo, mas que forneçam tais soluções em um tempo computacional compatível com o tempo geralmente concebido para projetos de engenharia.

Esta Tese possui como objetivo principal a investigação de novas técnicas para planejamento de redes ópticas WDM. A modelagem do projeto usa como dados de entrada: a posição geográfica dos nós, um algoritmo de RWA e uma demanda de tráfego. A proposta de planejamento de rede retornada pelo algoritmo de otimização deverá informar quais enlaces serão usados na rede, o número máximo de comprimentos de onda por enlace e o modelo de comutador óptico empregado. É imposta a restrição de que a topologia física encontrada seja um grafo conectado. A nova proposta deverá oferecer soluções de projeto de redes WDM com qualidade igual ou superior à apresentada por estudos anteriores, mas reduzindo o tempo total de planejamento para um tempo compatível com o uso prático do algoritmo para uma ferramenta de planejamento de redes.

Portanto, o objetivo geral desta Tese é investigar uma nova abordagem para planejamento de redes WDM, incluindo no estudo da Tese os seguintes aspectos:

1. Estudo e proposição de métricas que possam ser usadas para quantificação de características de topologia física de redes ópticas;
2. Estudo e proposição de procedimentos para criação de grafos que possam ser usados como modelos para geração de topologias físicas de redes ópticas;
3. Proposição de métodos alternativos para estimativa da PB de redes WDM que possam contribuir para a melhoria da proposta de planejamento;
4. Estudos de casos em redes conhecidas para permitir a avaliação da nova abordagem de planejamento de redes e comparação com trabalhos anteriores.

1.1 Organização do Documento

Este documento está estruturado em seis capítulos.

No Capítulo 2 é realizada uma revisão da literatura com relação aos temas centrais da tese. É feita uma revisão da literatura sobre trabalhos relacionados com análise de redes ópticas e

em seguida é feita uma revisão sobre os trabalhos relacionados com o planejamento de redes ópticas.

O Capítulo 3 apresenta a fundamentação teórica sobre a Ciência das Redes e as propostas da Tese relacionadas com essa área de investigação. São apresentadas métricas e modelos de geração de redes conhecidos na literatura. Em seguida, são apresentadas novas métricas que são propostas nesta Tese.

O Capítulo 4 apresenta detalhes sobre análise de redes ópticas por meio de simulação e apresenta propostas desta Tese sobre métodos alternativos de análise de redes. Neste capítulo é detalhada uma proposta de análise de redes baseada em técnicas de aprendizagem de máquinas e métricas topológicas de redes.

O Capítulo 5 versa sobre o planejamento de redes ópticas sob uma perspectiva de otimização multiobjetiva. Este capítulo inicia com um detalhamento sobre técnicas de otimização multiobjetivas e, em seguida, são apresentadas as novas propostas desta Tese para o planejamento de redes.

O Capítulo 6 faz um resumo das contribuições realizadas neste trabalho, apresenta as conclusões obtidas a partir desta Tese e apresenta sugestões para trabalhos futuros.

Capítulo 2

Revisão da Literatura em Análise e Projeto de Redes Ópticas

Sistemas de comunicações ópticas foram utilizados inicialmente para aumentar a taxa de transmissão ponto-a-ponto, visto que fibras ópticas oferecem uma largura de banda superior aos fios de cobre e evitam interferências eletromagnéticas e outros efeitos indesejáveis. Dessa forma, as primeiras implementações de sistemas de alta capacidade tipo WDM (*Wavelength Division Multiplexing*) foram do tipo ponto-a-ponto, com o objetivo de aumentar a capacidade de transmissão por fibra óptica em um enlace, aliviando a demanda pelo lançamento de novos cabos. Nestes primeiros sistemas, a capacidade de roteamento e demais serviços de inteligência da rede são desempenhados por dispositivos eletrônicos. As redes com essa arquitetura são conhecidas como redes ópticas de primeira geração, dentre as quais podem ser citadas as redes SONET (*Synchronous Optical Network*) e SDH (*Synchronous Digital Hierarchy*). A partir do desenvolvimento de amplificadores ópticos, de *lasers* de melhor qualidade e equipamentos mais precisos (filtros, multiplexadores e demultiplexadores), foi possível transmitir mais canais WDM simultaneamente, aumentando assim a capacidade total. Essa tecnologia ficou conhecida como tecnologia DWDM (*dense WDM*) [2].

Após o desenvolvimento de outros componentes, como OADMs (*optical add and drop multiplexers*) e ROADMs (*Reconfigurable optical add-drop multiplexers*), os sistemas ópticos têm deixado de ser apenas uma tecnologia de enlace de alta capacidade para se tornar uma tecnologia de rede de alta capacidade [23]. Essas novas arquiteturas oferecem serviços inteligentes da rede no domínio óptico, em oposição aos sistemas anteriores em que estes serviços são providos por dispositivos eletrônicos, com necessidade de conversão óptica-eletrônica-óptica (O-E-O). As redes que realizam conversão O-E-O em todos os nós intermediários entre um par origem-destino são conhecidas como **redes ópticas opacas**. As redes que não realizam conversão O-E-O nos nós intermediários são conhecidas como **redes totalmente ópticas**. As redes que realizam conversão O-E-O em alguns dos nós intermediários são conhecidas como

redes translúcidas [2].

Há também um crescente interesse em novas arquiteturas de redes ópticas, como por exemplo a denominada de *gridless*, que não possui uma grade fixa e uniformemente espaçada de comprimentos de onda, como ocorre no caso das tecnologias WDM ou DWDM. Em redes *gridless*, os equipamentos fornecem suporte para que a largura de banda dos caminhos ópticos seja flexível, permitindo expansão ou contração de espectro de acordo com a demanda [24]. A arquitetura *gridless* tem apresentado crescente interesse da comunidade acadêmica, mas ainda enfrenta desafios que impossibilitam o seu uso comercial em larga escala. O principal objeto de estudo desta Tese são redes ópticas WDM e DWDM e, portanto, as redes *gridless* não são mais discutidas daqui por diante. Entretanto, os conceitos e metodologias propostas e desenvolvidas nesta Tese podem ser adaptadas para redes do tipo *gridless*.

Neste Capítulo é feito um levantamento sobre trabalhos já realizados sobre análise e planejamento de redes ópticas, que são os dois principais problemas estudados nesta Tese. A Seção 2.1 trata dos trabalhos relacionados com análise de redes ópticas e a Seção 2.2 apresenta os trabalhos relacionados com o planejamento de redes ópticas.

2.1 Análise de Redes Ópticas

A análise de uma rede óptica pode ser realizada sobre diversos aspectos, tais como: análise econômica baseada em custos de implantação e de operação da rede, análise de consumo energético, análise de desempenho por meio de medidas de Qualidade de Transmissão (QoT), análise de tolerância a falhas, dentre outras.

Nesta Tese, o escopo do estudo será direcionado para a análise do desempenho de redes ópticas com restrições de QoT e análise do custo da rede. O modelo de custo que será adotado na Tese é o modelo de custo de implantação (CAPEX) usado em [21]. A Seção 2.1.1 detalha o modelo de custo adotado na Tese e as Seções 2.1.2 e 2.1.3 possuem um levantamento da literatura sobre análise de desempenho em enlaces e redes ópticas, respectivamente.

2.1.1 Modelo para Avaliação do Custo de uma Rede Óptica

Foi utilizado neste trabalho o mesmo modelo de custo adotado em [21], que é uma adaptação do modelo de custo de equipamentos ópticos proposto por Huelsermann *et al.* [25]. O modelo considera quatro diferentes fontes de custo: um custo fixo por comprimento de onda utilizado na rede inteira, o custo de implantação dos cabos de fibra óptica, o custo dos amplificadores ópticos e o custo dos dispositivos ROADMs. Uma unidade monetária genérica (*u.m.*) foi utilizada para comparar o custo das diversas configurações de rede.

Cada comprimento de onda da rede está associado a um custo devido a dois *transponders* OLT - *Optical Line Terminal*, ou seja, um dispositivo para conversão óptico eletrônico

que ficará entre os clientes e a rede óptica transparente. O custo associado ao número de comprimentos de onda é definido por

$$COST_{Lambda} = 2 \cdot \eta \cdot W \cdot \sum_{i=1}^n \mathcal{D}(i, i), \quad (2.1)$$

em que W é o número de comprimentos de onda por enlace, η é um valor constante que pode ser inferido por meio do custo do equipamento OLT, n é o número de nós da rede e $\mathcal{D}(i, i)$ é o grau do i -ésimo nó. O grau de um nó corresponde ao número de enlaces que partem ou que chegam no nó.

A rede possui um conjunto de enlaces conectando os nós da rede. Existem dois custos envolvidos neste processo, o custo do cabo de fibra óptica e o custo de implantação da fibra. A Eq. (2.2) é utilizada para calcular o custo em *u.m.* por quilômetro.

$$COST_{Cable} = (\beta_{dcf} + \beta_{ssmf} + \beta_{Dep}) \sum_{i=1}^n \sum_{j=i+1}^n d_{i,j}, \quad (2.2)$$

em que $d_{i,j}$ é a distância física entre os nós i e j . β_{dcf} e β_{ssmf} são os valores de entrada do custo da fibra DCF - *Dispersion Compensate Fiber* e da fibra SSMF - *Standard Single Mode Fiber*. β_{Dep} é um valor de entrada relacionado ao custo de implantação da fibra.

Os custos dos amplificadores ópticos são considerados de acordo com a potência de saturação na saída e a figura de ruído. O custo total relacionado aos amplificadores é dado por

$$COST_{Amplifier} = \delta \cdot \sum_{i=1}^{e_{max}} C_{amp}(i), \quad (2.3)$$

em que $e_{max} = \frac{n(n-1)}{2} = \frac{n^2-n}{2}$, $C_{amp}(i)$ é o custo unitário de um modelo de amplificador previamente rotulado e δ é uma constante de custo para tornar o modelo genérico.

Um método semelhante é usado para definição do custo do comutador óptico ($COST_{ROADM}$). O custo total dos ROADMs é dado por

$$COST_{ROADM} = C_{sw}(v_{K+1}) \cdot \sum_{i=1}^n ((\gamma_1 \cdot W + \gamma_2) \cdot \mathcal{D}(i, i) + \gamma_3), \quad (2.4)$$

em que $\mathcal{D}(i, i)$ é o grau do i -ésimo nó e $C_{sw}(\ell)$ está relacionado ao fator de isolamento do ROADM. Este modelo possui como premissa o uso do mesmo ROADM em todos os nós da rede. γ_1 , γ_2 , γ_3 são constantes de custo que podem ser adaptadas para adequar o modelo às condições de mercado.

Considerando as quatro diferentes fontes de custo, o custo total da rede ($COST_{Net}$) é definido por

$$COST_{Net} = COST_{Lambda} + COST_{Amplifier} + COST_{Cable} + COST_{ROADM}. \quad (2.5)$$

2.1.2 Análise da Qualidade do Sinal em Enlaces Ópticos

A análise da degradação do sinal em um enlace óptico pode ser feita por meio de três principais estratégias: uso de cálculos numéricos de propagação de onda em fibras ópticas e em dispositivos ópticos (abordagem numérica) [5, 6]; uso de medição direta do sinal em propagação na rede (abordagem experimental) [26, 27]; uso de expressões analíticas aproximadas que estimam as penalidades da camada óptica [28, 29, 30, 2].

O uso da abordagem numérica é muito precisa, contudo o custo computacional é elevado e restringe o uso deste método para cenários mais simples. A abordagem experimental também é precisa, mas exige que o enlace já tenha sido implantado. Portanto, a abordagem experimental é inadequada para a fase de planejamento de um novo enlace. Considerando o elevado risco em se implantar um enlace óptico sem um estudo prévio de desempenho, é interessante possuir um mecanismo de análise que possa ser usado na fase de projeto do enlace e ofereça um custo computacional baixo. Um bom compromisso entre precisão e custo computacional é obtido por meio do uso de expressões analíticas simplificadas que fornecem uma estimativa da degradação do sinal.

Esta Tese usa simulações para análise e planejamento de redes ópticas, mas estas simulações são baseadas em expressões analíticas para estimativa da degradação do sinal de um caminho óptico individual. Nas subseções seguintes será feita uma explanação sobre alguns indicadores que podem ser usados para quantificar a degradação do sinal óptico. Na Seção 2.1.3 será explicada a relação destes indicadores com a análise global da rede óptica.

Relação Sinal Ruído

A medida em que um sinal se propaga pelo caminho óptico entre o transmissor e o receptor, diversas fontes podem inserir ruído sobre o sinal transmitido. O ruído inserido pode ser severo o bastante para impedir que o receptor consiga decodificar corretamente o sinal recebido. Portanto, a relação sinal ruído óptica (OSNR - *Optical Signal to Noise Ratio*) é uma importante medida a ser monitorada em um caminho óptico. Quanto maior a OSNR, melhor a qualidade do sinal. A degradação da OSNR em redes ópticas transparentes é devida principalmente ao acúmulo de ruído de emissão espontânea amplificada (ASE), *crosstalk* e efeitos não-lineares. Para calcular a OSNR, os dois parâmetros que são medidos (normalmente nos receptores) são: a potência média do sinal óptico recebido (P_S) e a potência do ruído óptico recebido (P_N). A OSNR pode ser calculada, em decibéis, usando a Eq. (2.6) [2].

$$\text{OSNR} = 10 \cdot \log_{10} \left(\frac{P_S}{P_N} \right). \quad (2.6)$$

Fator de Ruído e Figura de Ruído

O fator de ruído (F - *Noise Factor*) é usado para quantificar a degradação da relação sinal ruído devido ao ruído inserido pelos amplificadores e pelos diversos componentes ópticos dispostos ao longo do caminho óptico. Em termos quantitativos, se OSNR_{in} é a relação sinal ruído na entrada do dispositivo e OSNR_{out} é a relação sinal ruído na saída do dispositivo, ambas em escala linear, o fator de ruído é definido como sendo a razão entre a OSNR_{in} e a OSNR_{out} , conforme a Eq.(2.7) [2]:

$$F = \frac{\text{OSNR}_{in}}{\text{OSNR}_{out}}. \quad (2.7)$$

A figura de ruído (NF - *Noise Figure*) é o fator de ruído expresso em unidades de decibéis (dB) e é dado pela Eq.(2.8) [2]:

$$NF = 10 \cdot \log_{10} F. \quad (2.8)$$

Taxa de Erro de Bit do Sinal

A taxa de erro de bit (BER - *Bit Error Rate*) quantifica o desempenho da transmissão pelo meio físico e também é utilizada como uma medida de QoT. Em um sistema de comunicação digital com modulação OOK (*On-off keying*) em que o sinal recebido representa um feixe de 0s e 1s, é preciso decidir na chegada de um pulso óptico qual a sequência correta de 0s e 1s foi recebida. Para decidir, é necessário estabelecer um nível limite para o sinal recebido. Sempre que o sinal recebido estiver acima desse nível, o bit será considerado como bit 1; e quando estiver abaixo do nível, será considerado como sendo o bit 0. Se o sinal recebido não tiver potência suficiente no momento da detecção, devido às perdas (atenuação), ou se a quantidade de ruído inserido no sinal for elevada, poderá ocorrer erros na tomada de decisão. Para o sistema binário, o receptor poderia detectar um valor 0 ao invés de 1 ou vice-versa. O mesmo conceito se aplica a outros formatos de modulação digital. Uma BER de $2 \cdot 10^{-6}$ significa 2 bits errados em 10^6 bits transmitidos.

Alargamento Temporal do Pulso Óptico

A transmissão em fibra óptica também está sujeita a alguns efeitos de dispersão que alargam o pulso óptico e podem afetar a recepção do sinal devido à interferência inter-simbólica. Os efeitos mais comuns que estão relacionados com o alargamento do pulso são: dispersão cromática e dispersão por modo de polarização (*Polarization-Mode Dispersion* - PMD) [2].

A dispersão cromática é o fenômeno físico relacionado com o fato de que diferentes componentes espectrais de um pulso percorrem um caminho óptico a velocidades diferentes. A maior parcela de alargamento temporal devido à dispersão cromática está associada ao fato de que o índice de refração da sílica é dependente da frequência considerada no sinal óptico. Além disso, fibras ópticas reais não são perfeitamente cilíndricas, de modo que os dois modos polarizados ortogonalmente possuem constantes de propagação ligeiramente diferentes. O alargamento do pulso óptico devido a esta característica é conhecido como dispersão por modo de polarização (*Polarization-Mode Dispersion* - PMD) [2].

Um alargamento temporal de $\Delta\tau = 0,5$ ps/km, por exemplo, significa que após a propagação de 100 km em uma fibra óptica, o alargamento temporal será de 50 ps [2]. Um valor de limiar de alargamento temporal pode ser usado no receptor para identificar se o alargamento do pulso é aceitável do ponto de vista de QoT.

2.1.3 Análise de Desempenho em Redes Ópticas

A análise de desempenho em uma rede óptica pode ser feita por meio do uso de diversas métricas, dentre as quais podem ser citadas: vazão (*throughput*) [31], atraso médio [32] e probabilidade de bloqueio [33].

A probabilidade de bloqueio (PB ou BP, do inglês *Blocking Probability*) é um indicador que tem sido usado em diversos estudos para medir o desempenho de redes ópticas sujeitas a tráfego dinâmico. Nesta Tese a PB será adotada como medida de desempenho de rede e, portanto, serão fornecidos mais detalhes sobre esta métrica.

PB é uma medida estatística referente às requisições de chamadas que não puderam ser estabelecidas devido à falta de recursos da rede. É uma medida de desempenho global da rede. Quanto mais elevado for o valor da probabilidade de bloqueio, menos usuários serão atendidos e, conseqüentemente, pior será o desempenho da mesma. Uma estratégia comum para obter uma estimativa de PB é simular um grande número de requisições de chamadas à rede óptica e contabilizar a quantidade de chamadas atendidas e a quantidade de chamadas negadas. Usando este método, a probabilidade de bloqueio pode ser estimada por

$$PB = \frac{C_{bloqueadas}}{C_{total}}, \quad (2.9)$$

em que $C_{bloqueadas}$ é a quantidade de chamadas bloqueadas e C_{total} é a quantidade total de chamadas.

A parcela de chamadas bloqueadas pode ser influenciada devido à ausência de QoT avaliada de acordo com um ou vários dos indicadores apresentados nas seções precedentes. Ou seja, é estabelecido um valor de limiar para cada métrica e este valor é comparado com o valor obtido no receptor. Por exemplo, seja $OSNR_{TH}$ um requisito de OSNR de limiar para avaliação de QoT em um caminho óptico. Se $OSNR_{receptor} < OSNR_{TH}$ no momento da ava-

liação de uma dentre as várias requisições usadas para estimativa de PB, esta chamada será bloqueada pelo critério de QoT insuficiente devido à baixa relação sinal ruído óptica. Os outros indicadores podem ser considerados de forma análoga.

Considerando que a estimativa da PB depende da análise dos indicadores de qualidade apresentados na Seção 2.1.2, em princípio seria possível usar quaisquer das estratégias já definidas na Seção 2.1.2 para computar a parcela de chamadas bloqueadas. Contudo, o elevado custo computacional da abordagem numérica torna proibitivo o uso desta estratégia para um cálculo preciso de PB, ficando sua aplicação restrita à análise de enlaces ópticos. Portanto, um método comum para estimar PB é por meio de simulações que usam expressões analíticas para calcular a qualidade do sinal óptico. Esta estratégia é adotada no simulador SIMTON [11] e também será usada nesta Tese. Além do uso de simulações para estimativa de PB, outros estudos tentaram obter expressões analíticas fechadas para a PB de uma rede óptica [33, 34, 35, 36]. Em geral, é muito difícil obter uma estimativa precisa de PB usando apenas expressões analíticas fechadas, pois esta medida está intimamente relacionada com a dinâmica da rede, incluindo o modelo e carga de tráfego e o número de comprimentos de onda disponíveis em cada enlace. As expressões analíticas propostas para PB geralmente ignoram as penalidades da camada física, consideram um número fixo de comprimentos de onda nos enlaces ou estabelecem premissas rígidas para a topologia física. Em geral, estes métodos são adequados apenas para redes muito regulares, como redes lineares ou em anel.

Diversas ferramentas de análise baseadas em simulação já foram propostas e apresentam várias diferenças entre si. Embora cada simulador proposto possua seu próprio elenco de recursos, é possível diferenciar as ferramentas em dois grandes grupos. No primeiro grupo estão as ferramentas que não consideram a degradação do sinal devido aos efeitos da camada física e são aplicados somente a redes ópticas opacas. Este grupo inclui: o simulador OWNS [9] (que é uma extensão do Network Simulator [37]); OPNET [7] e TONetS [10]. O segundo grupo de ferramentas de simulação considera a degradação do sinal devido às penalidades da camada física e, portanto, podem ser usadas para simulação de redes transparentes e translúcidas. Neste segundo grupo, enquadram-se as ferramentas: SIMON [8] e SIMTON [11].

Alguns estudos já demonstraram que é possível aplicar técnicas de aprendizagem de máquinas para diminuir o custo computacional de ferramentas para análise de enlaces ópticos. Jargon *et. al* [38] desenvolveram um modelo de rede neural artificial (RNA) para identificar separadamente três penalidades da camada física (OSNR, dispersão cromática e PMD). Neste trabalho a RNA foi treinada a partir de parâmetros de diagramas de olho. Shen *et. al* [39] desenvolveram um modelo de RNA semelhante, mas treinado a partir de histogramas de amostragem assíncrona (*Asynchronous Amplitude Histograms* - AAHs). Apesar do uso de uma RNA para as aplicações de monitoramento do sinal óptico terem demonstrado serem promissoras, não foram encontrados trabalhos que usam técnicas de aprendizagem de máquinas para predição de um indicador de desempenho global de rede.

2.2 Planejamento de Redes Ópticas

O planejamento de redes ópticas WDM pode ser classificado em dois grupos: o projeto de topologia virtual ou lógica (*Virtual Topology Design* - VTD) [40] e o projeto de topologia física (*Physical Topology Design* - PTD) [14]. A topologia física estabelece a ligação entre nós de uma rede por meio de fibra óptica. Ela provê a infraestrutura de telecomunicações necessária para a comunicação na rede óptica. Por outro lado, a topologia lógica estabelece os circuitos virtuais (ou caminhos ópticos físicos) conectando os diversos pares de nós. A topologia lógica precisa de uma topologia física subjacente, mas as duas podem ser distintas [2]. O problema de VTD se assemelha ao problema de roteamento e atribuição de comprimento de onda quando um tráfego estático é considerado. Neste trabalho será considerado o problema de planejamento de topologia física e o problema VTD não será mais discutido. Há uma vasta literatura sobre VTD para o leitor que deseja se aprofundar [40].

O PTD é classificado como um problema NP-completo [14] ou NP-difícil [13], dependendo da forma como o problema é modelado e dos requisitos necessários para a rede que é planejada. Ambos os tipos de problemas são computacionalmente complexos, pois requerem uma quantidade de tempo impraticável para se obter uma solução exata do problema em um computador convencional. Para PTD, esta complexidade cresce exponencialmente de acordo com o número de nós da rede planejada. Os trabalhos que se propõem a obter a solução ótima deste tipo de problema usam apenas redes muito pequenas [17]. Para realizar o planejamento de redes reais, que geralmente possuem até dezenas de nós, é necessário o uso de técnicas que forneçam soluções subótimas, mas em um tempo aceitável [15, 41, 16]. Estas técnicas podem ser heurísticas [14, 41, 16] ou meta-heurísticas [13, 17, 32, 42, 43, 44, 45].

Heurísticas são técnicas que usam alguma informação específica do problema para obter soluções aproximadas. Em trabalhos que usam soluções heurísticas é muito comum encontrar variações do algoritmo *branch exchange* (BE) [46] e *cut saturation* (CS) [47]. *Branch exchange* consiste em criar uma rede totalmente conectada e realizar remoções de enlaces iterativamente até alcançar um efeito desejado [46]. *Cut saturation* inicia o processo com uma árvore mínima e acrescenta ou remove enlaces usando o critério de corte saturado (*saturated cut*) [47]. Meta-heurísticas são algoritmos genéricos criados para serem usados em uma determinada classe de problemas, precisando apenas que o projetista realize pequenas modificações para aplicação em um problema específico. São exemplos de uso de meta-heurísticas para PTD: trabalhos que usam *simulated annealing* [14, 41, 16, 45] trabalhos que usam *tabu search* e trabalhos que usam algoritmos genéticos [13, 17, 32, 42, 43, 44, 48].

Apesar do problema de PTD ter sido estudado por muitos anos, trabalhos específicos para o planejamento de redes ópticas são mais recentes. Dos trabalhos já realizados com foco em planejamento de redes ópticas, é importante ressaltar algumas diferenças. Muitos trabalhos focaram no estudo de planejamento de redes ópticas sujeitas a tráfego estático [14, 15, 16,

17, 18, 19, 20]. Além disso, estes trabalhos modelaram o problema de planejamento de uma rede óptica com o foco na minimização de um objetivo único, como o custo da rede [14, 48] ou o número total de comprimentos de onda a serem usados [18, 16]. O modelamento do problema com objetivo único ignora o fato conhecido de que projetos reais possuem objetivos conflitantes, como custo *versus* desempenho da rede, consumo energético *versus* desempenho, dentre outros. Os trabalhos encontrados na literatura com uma modelagem que trata mais de um objetivo conflitante é mais escassa. Nesta categoria podem ser citados o trabalho sobre otimização multiobjetiva em redes eletrônicas de Banerjee e Kumar [32] e em redes ópticas de Chaves *et. al* [49] e Araujo *et. al* [22].

À medida que a modelagem do problema PTD para redes ópticas considera dois ou mais objetivos conflitantes, e usa para análise de desempenho da rede alguma ferramenta que considera os efeitos da camada física, o algoritmo de planejamento fornece soluções mais confiáveis, pois considera aspectos do problema que estavam sendo negligenciados por abordagens mais simples. Contudo, o custo computacional para execução destes algoritmos de planejamento de topologia física torna-se muito elevado e em alguns casos o planejamento de uma rede óptica nestas condições pode demorar diversos dias [49, 22].

Por outro lado, estudos realizados para redes americanas e europeias indicam que existe uma forte relação entre aspectos econômicos e demográficos, e a malha de fibra óptica que é implantada para atender as demandas por tráfego [50, 51, 52]. Moss e Townsend [50] estudaram o desenvolvimento da Internet nos Estados Unidos da América no período de 1977 até 1999 e estabeleceram que há uma forte correlação entre a topologia física e lógica das redes de transporte com a presença de empresas que possuem a informação como principal ativo. Recentemente, Tranos e Gillespie [51] estudaram os fatores que determinam a distribuição espacial das redes de transporte na Europa e concluíram que diversas variáveis afetam a topologia física mas, em geral, os nós apresentam maior conectividade em grandes regiões metropolitanas. Tranos [52] comparou a infraestrutura de redes de aviação com a da Internet na Europa e concluiu que a infraestrutura aérea possui uma distribuição em lei de potência, mas as redes de transporte ópticas são mais estruturadas. Cardenas *et. al* [53] apresentaram um estudo em que mostram que a distribuição de grau dos nós de redes SDH operadas pela Telefônica-Espanha também é caracterizada por lei de potência. Um estudo sobre a emergência de modelos canônicos de redes em topologias de redes da Internet de diversos países foi realizado por Knight *et. al* [54]. Este último estudo não foi conclusivo sobre quais modelos canônicos são mais adequados para representar uma rede de transporte, pois algumas redes apresentam distribuição em lei de potência (*power-law*) e outras são mais regulares (não demonstram possuir concentradores). Apesar de existirem trabalhos que tentam correlacionar a topologia de redes de transporte ópticas com métricas topológicas e modelos de Ciências das Redes, há poucos estudos na literatura que usam este tipo de conhecimento como parte integrante de ferramentas para planejamento de novas redes.

Dharmaweera *et. al* [55] propuseram uma heurística para planejar redes ópticas considerando consumo de energia, custo e resiliência. Neste estudo foram usadas duas redes existentes na Austrália para avaliação da heurística proposta [55]. Paiva *et. al* [56] propuseram o uso de *twin graphs* para o planejamento de redes OTN. Neste estudo o foco foi na facilidade inerente ao modelo *twin graphs* para a proposição de topologias físicas com boas características de proteção contra falhas [56]. Cetinkaya *et. al* [57] fizeram uma análise comparativa sobre o uso de 4 modelos de geração de grafos para explicar a topologia física de redes de transporte implantadas nos Estados Unidos. Neste estudo, as topologias criadas foram comparadas com topologias de redes reais usando inspeção visual dos grafos da rede e métricas espectrais. Um modelo de custo que considera o comprimento dos enlaces de fibra óptica foi usado para comparar o custo das redes geradas e das redes reais [57]. Maniadakis e Varoutas [58] aplicaram grafos de Gabriel em redes de acesso do tipo *fiber-to-the-x* (FTTx) e identificaram que este modelo reproduz algumas métricas topológicas encontradas na malha de fibra das ruas das cidades. Contudo, nenhum destes estudos incluiu diretrizes para planejamento de redes e aspectos relacionados com restrições de QoT para identificar se as topologias de rede são adequadas para uso em situações reais nas quais a qualidade do sinal seja uma restrição de projeto.

Esta Tese propõe uma abordagem de planejamento baseada nos avanços obtidos em Ciência das Redes e em Aprendizagem de Máquinas e avalia as redes propostas pelo algoritmo de planejamento por meio de estimativa de PB influenciada por diversos critérios de QoT. No Capítulo 3, são apresentados conceitos e métricas de Ciência das Redes e também são propostas novas métricas que podem fornecer informações relevantes sobre redes ópticas. No Capítulo 4 é proposta uma nova abordagem para estimativa de PB baseada em Aprendizagem de Máquinas. O Capítulo 5 fornece o novo algoritmo de otimização para planejamento de redes ópticas.

Capítulo 3

Ciência das Redes

Ciência das redes é uma área de investigação interdisciplinar que estuda redes complexas presentes em diferentes domínios de aplicação, tais como as redes de telecomunicações, redes biológicas, redes neurais, redes sociais, dentre outras. O *National Research Council* define ciência das redes como o estudo das representações de rede de fenômenos físicos, biológicos e sociais, levando a modelos preditivos desses fenômenos [59].

Avanços importantes na Ciência das Redes estão relacionados com a proposição de modelos de geração de redes para criar grafos com propriedades topológicas semelhantes às propriedades encontradas em redes do mundo real. Em 1960, Erdos e Renyi apresentaram estudos importantes relacionados a redes aleatórias [60]. Watts e Strogatz propuseram, em 1988, o primeiro procedimento para geração de redes que apresentam o efeito de mundo pequeno (*small-world* - SW) [61]. Em 1999, Barabási e Albert apresentaram um modelo baseado em conexão preferencial (*preferential attachment*) para gerar redes livres de escala (*scale-free* - SF) [62]. Diversas variantes destes modelos já foram propostas nos últimos anos. Uma revisão completa destes modelos pode ser obtida em [63].

Na maior parte dos casos, as redes reais não apresentam uma topologia com características aleatórias, como as redes geradas pelo modelo Erdos-Renyi (ER). Em geral, as redes reais apresentam características mais próximas de redes regulares, redes com efeito de mundo pequeno, redes livre de escala ou uma composição destas características. Portanto, as características topológicas podem ser usadas para classificar as redes em famílias de grafos. Este aspecto é muito importante para a compreensão da relação entre características topológicas e o comportamento das redes reais. Por exemplo, considere a análise de famílias de redes por meio de métricas topológicas de robustez, tais como conectividade algébrica ou natural, comprimento médio dos caminhos e tamanho do componente mais conectado. Neste tipo de análise é possível descobrir que as redes livres de escala, criadas pelo modelo Barabási-Albert (BA), estão relacionadas com redes reais que apresentam elevada resiliência a falhas em nós aleatórios, mas são muito vulneráveis a ataques direcionados. Por outro lado, redes aleatórias

são robustas a ataques direcionados [64]. O planejamento de novas redes que serão usadas em uma aplicação real envolve em primeira instância a escolha de uma topologia física que promova padrões específicos para o comportamento dinâmico da rede. Em geral, estes padrões estão associados com requisitos de projeto para a rede. Portanto, um elemento chave em planejamento de redes é conseguir identificar se uma topologia física segue um determinado modelo teórico de redes, pois os modelos teóricos já foram exaustivamente estudados e possuem características conhecidas. O estudo de características topológicas para prever a aderência da rede com os requisitos de projeto é uma área de investigação recente [65]. Contudo, para permitir a correta correlação entre métricas topológicas e requisitos da rede, é necessário dispor de um conjunto de métricas expressivas e que sejam sensíveis a pequenas alterações na topologia física.

Neste capítulo serão apresentadas as contribuições da tese com relação a métricas topológicas para análise de redes. A Seção 3.1 fornece um resumo de características topológicas conhecidas na literatura. A Seção 3.2 apresenta os modelos de geração de redes mais conhecidos e estabelece uma relação entre as redes geradas pelos modelos e as métricas topológicas estudadas. As seções 3.3, 3.4 e 3.5 são contribuições desta tese para a análise de redes por meio de métricas topológicas.

3.1 Características Topológicas de Redes

Uma rede pode ser modelada como um grafo $G = (\mathcal{N}, \mathcal{E})$, em que \mathcal{N} é um conjunto rotulado de nós ou vértices e \mathcal{E} é um conjunto de enlaces ou arestas que representam a ligação entre os nós da rede. Para quantificar o número de nós e o número de enlaces da rede, usam-se $n = |\mathcal{N}|$ e $e = |\mathcal{E}|$, respectivamente. O relacionamento entre os nós da rede pode ser estudado por meio de métricas que são calculadas em função dos elementos do conjunto \mathcal{E} . Um formato típico de representação de G é por meio de matrizes de adjacências. Uma matriz de adjacências \mathcal{A} indica a ligação entre dois nós i e j de G , se um elemento $a_{i,j} = 1$, e indica ausência de conectividade se $a_{i,j} = 0$. Se a rede representada por G possui apenas conexões bidirecionais, a matriz \mathcal{A} é simétrica.

Um grafo é dito conectado quando existe um caminho entre todos os pares de nós origem-destino. Quando não há caminho para conectar pelo menos um par de nós da rede, o grafo é dito desconectado. A conectividade por enlace (k_l) é o número mínimo de enlaces que precisam ser removidos de modo a transformar um grafo conectado em um grafo desconectado. A conectividade por nós é definida de forma análoga [66]. A **densidade** de uma rede (q) é definida como a taxa entre o número de enlaces que existem na rede e o número máximo de enlaces que poderiam existir. Portanto,

$$q = \frac{e}{n(n-1)}. \quad (3.1)$$

O **grau de um nó** (d) representa o número de enlaces que ligam um nó aos seus nós vizinhos. O número médio de enlaces que estão conectados a um nó é denominado **grau médio** e é definido por

$$\hat{d} = \frac{1}{n} \sum_{i=1}^n d_i. \quad (3.2)$$

A **distribuição de grau dos nós** de G (g') estima a probabilidade $Pr(d)$ de que um nó selecionado ao acaso, possua um certo grau d .

O **coeficiente de assortatividade** de uma rede ($-1 \leq r \leq 1$) avalia se há uma tendência de que os nós se conectem com nós que possuem grau semelhante ($r \geq 0$). Se $r \leq 0$, há uma tendência de que os nós com grau elevado se conectem com nós que possuem grau baixo. O coeficiente de assortatividade r pode ser obtido por meio de

$$r = \frac{\sum_{i,j \in E} d_i d_j - \sum_{i,j \in E} (d_i + d_j)^2 / 2e}{\sum_{i,j \in E} (d_i^2 + d_j^2) - \sum_{i,j \in E} (d_i + d_j)^2 / 2e}. \quad (3.3)$$

O **caminho mais curto** ($SP_{i,j}$) descreve o número mínimo de saltos necessário para caminhar entre dois pares de nós i e j . O maior valor de caminho mais curto na rede, quando são considerados todos os pares origem-destino, é conhecido como **diâmetro da rede** ($diam(G) = \max(SP_{i,j}, \forall i, j)$). O **comprimento médio dos caminhos** mínimos (\bar{c}) é a média de todos os SP entre todos os possíveis pares origem-destino. A **conectividade natural** de um grafo (NC) é uma medida do número de cadeias fechadas de quaisquer comprimentos [67]. Ou seja, NC é uma medida de robustez, pois fornece a informação de caminhos alternativos na rede.

O coeficiente de agrupamento de um nó i (cc_i) é a razão entre o número de triangulações que contém o nó i e o número de triangulações que poderiam existir se todos os nós vizinhos de i estivessem conectados. Considerando um nó i , existem d_i vizinhos e potencialmente $d_i(d_i - 1)/2$ enlaces conectando todos os vizinhos. Supondo que estes vizinhos compartilham c enlaces, pode-se calcular cc_i por meio de

$$cc_i = \frac{2c}{d_i(d_i - 1)}. \quad (3.4)$$

O **coeficiente de agrupamento** da rede (CC) é a média de todos os possíveis c_i da rede, ou seja,

$$CC = \frac{1}{n} \sum_{i=1}^n cc_i. \quad (3.5)$$

Uma rede com elevado coeficiente de agrupamento indica que há alguns grupos de nós na rede que são altamente conectados entre si, mas a ligação entre estes grupos é feita por meio

de poucos enlaces. Considere o cálculo de CC para a rede de 10 nós ilustrada na Figura 3.1. O cálculo de cc para o nó i é dado por $cc_i = (2 \cdot 1)/[3(3 - 1)] = 2/6 = 0,33$, pois há apenas um enlace que é compartilhado pelos vizinhos do nó i (enlace $\{k, l\}$) e a quantidade máxima de enlaces neste grupo de nós é $3(3 - 1) = 6$. Se o mesmo procedimento for repetido para todos os 10 nós, encontra-se $CC = 0,42$.

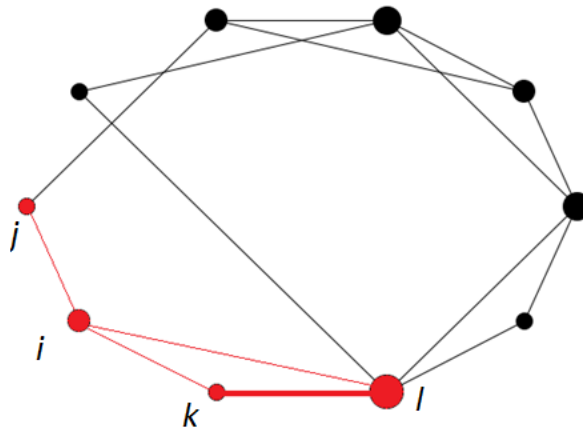


Figura 3.1: Exemplo de cálculo da métrica CC para uma rede de 10 nós.

Algumas métricas também foram definidas para avaliação de aspectos de centralidade das redes. Estas métricas tentam quantificar a importância de um determinado nó ou enlace perante a rede como um todo. Um mecanismo básico para avaliação da centralidade de um nó é contabilizar o número de arestas incidentes no nó. Ou seja, o grau de um nó é a métrica de centralidade mais simples conhecida. Outra métrica de centralidade considera o conceito de intermediação (*betweenness*) e avalia a quantidade de caminhos mínimos que passam por um determinado nó ou enlace da rede. Os nós (ou enlaces) que são usados para compor muitos caminhos mínimos (*SPs*) tendem a apresentar valor elevado de *betweenness* [68]. Outra medida de centralidade de nós amplamente usada é a centralidade por proximidade (*closeness*), que avalia a distância média de um a todos os demais nós da rede, considerando novamente os caminhos mínimos (*SPs*) [63]. Além disso, também é útil quantificar o capacidade de um grafo em facilitar o fluxo de informações por meio de métricas de centralidade que consideram a rede como um todo. Freeman [1] realizou um estudo abrangente sobre centralidade de grafos e propôs a métrica de centralização, que é dada por

$$C_x = \frac{\sum_{i=1}^n [C_x(p_*) - C_x(p_i)]}{\max \sum_{i=1}^n [C_x(p_*) - C_x(p_i)]} \quad (3.6)$$

em que $C_x(p_i)$ é uma medida de centralidade de um nó i , $C_x(p_*)$ é a maior medida de centralidade na rede e $\max \sum_{i=1}^n [C_x(p_*) - C_x(p_i)]$ é a maior soma das diferenças do ponto central C_x para qualquer grafo com o mesmo número de nós da rede. Perceba que qualquer medida de centralidade de nó (como grau dos nós, *betweenness* e *closeness*) pode ser usada

para o cálculo de centralização. Freeman analisou a centralização aplicada a diversas métricas de centralidade e oferece o valor de $\max \sum_{i=1}^n [C_x(p_*) - C_x(p_i)]$ para cada caso [1].

A distribuição de grau dos nós $Pr(d)$ mede a probabilidade de que um nó da rede, escolhido ao acaso, possua o grau d . O nível de aleatoriedade de uma rede pode ser quantificado por meio da **entropia** da rede $I(G)$. Se a distribuição de grau dos nós de uma rede for dada por $Pr(d) = \{h_1, h_2, \dots, h_m\}$, em que h_i é a fração de nós da rede que apresentam grau i , $I(G)$ é calculada por meio de

$$I(G) = - \sum_{i=1}^m h_i [\log_2(h_i)]. \quad (3.7)$$

Outro grupo de métricas usado para análise de redes é derivado dos autovalores de matrizes que representam a rede. Além da matriz de adjacências (\mathcal{A}), são usadas na análise espectral a matriz de grau (\mathcal{D}) e a matriz Laplaciana da rede (\mathcal{L}). \mathcal{D} é uma matriz diagonal $n \times n$ que contém o grau dos nós na diagonal principal da matriz. A matriz \mathcal{L} é definida como $\mathcal{L} = \mathcal{D} - \mathcal{A}$. O conjunto ordenado de n autovalores de \mathcal{A} ou de \mathcal{L} é conhecido como espectro de G . Se existem dois grafos com conjuntos similares de autovalores, é provável que estes dois grafos apresentem uma estrutura semelhante ou um isomorfismo [69]. O maior autovalor de \mathcal{A} é denominado **raio espectral** (ρ). Esta métrica é usada para análise de fluxo em redes, como em estudos sobre disseminação de doenças [70]. O segundo menor autovalor de \mathcal{L} é denominado **conectividade algébrica** (λ_{n-1}). Um grafo é desconectado se $\lambda_{n-1} = 0$. Além disso, se $\lambda_{n-i+1} = 0$ e $\lambda_{n-i} \neq 0$ então o grafo G possui exatamente i componentes. Ou seja, o número de valores iguais a zero no conjunto de autovalores de \mathcal{L} fornece a informação sobre o número de componentes do grafo. A conectividade algébrica também fornece uma medida acerca da robustez com relação à conectividade da rede. Um elevado valor de λ_{n-1} indica uma maior dificuldade para dividir o grafo em duas componentes quando arestas são removidas deste grafo [71]. A função de densidade dos autovalores $f_\lambda(t)$ muitas vezes é usada quando o conjunto de autovalores possui muitos elementos, ou seja, para elevados valores de n . A função de densidade dos autovalores é obtida por

$$f_\lambda(t) = \frac{1}{n} \sum_{m=1}^n \delta(t - \lambda_m), \quad (3.8)$$

em que $\delta(t)$ é a função de Dirac [72].

As métricas apresentadas nesta seção são muito usadas em diversos estudos para indicar propriedades de desempenho esperadas em redes reais e qual modelo canônico melhor representa estas redes.

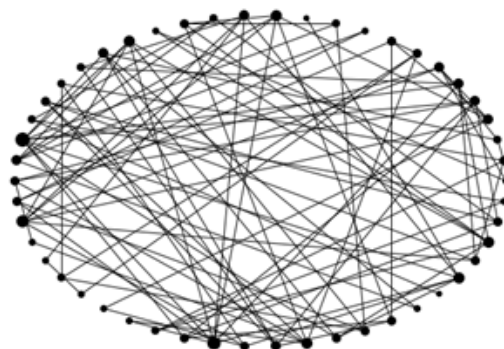
3.2 Modelos Canônicos de Redes

As redes aleatórias foram as primeiras redes exaustivamente estudadas. Apesar de as redes encontradas no mundo real sempre apresentarem algum padrão de regularidade, as redes aleatórias oferecem uma opção de comparação quanto ao grau e tipo de regularidade. As redes aleatórias possuem um valor alto de entropia e o coeficiente de agrupamento apresenta um valor baixo quando comparado com os valores apresentados pelas redes regulares. A assortatividade das redes aleatórias é próxima de zero. Os dois modelos mais conhecidos para gerar redes aleatórias são os modelos de Erdos-Renyi (ER) e Gilbert. O modelo ER gera redes por meio da inclusão de enlaces aleatórios na rede, até que uma densidade alvo seja alcançada. O método de Gilbert parte de uma rede completamente conectada e remove arestas escolhidas aleatoriamente, até que uma densidade alvo seja alcançada. $Pr(d)$ segue uma distribuição de Poisson para n grande e uma binomial para n pequeno [63].

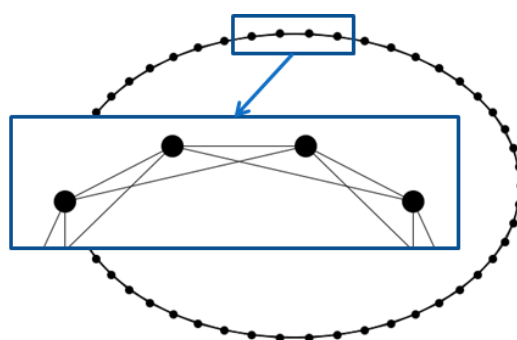
Redes k -Regulares são formadas por meio da ligação de cada nó i com os nós $\{i+1, \dots, i+k\}$, para $k \geq 2$. Se $k = 1$ a rede recebe o nome especial de rede em anel. As redes k -regulares são muito estruturadas, pois todos os nós possuem grau $d = 2k$. Logo, a entropia do grau dos nós de uma rede k -regular é zero. Uma rede k -regular é esparsa se k é pequeno e é densa se k é grande. A densidade pode ser obtida usando a expressão $\frac{2k}{n-1}$. Além disso, os grafos regulares são conectados e possuem o diâmetro e o comprimento médio dos caminhos proporcionalmente reduzidos quando k cresce [63]. A assortatividade das redes regulares é igual a um. Estas propriedades conferem ao modelo k -regular o papel de candidato natural para o projeto de redes reais. $Pr(d)$ é uma função delta, pois todos os nós possuem o mesmo grau.

As redes que possuem o efeito de mundo pequeno ou *small-world* apresentam um elevado coeficiente de agrupamento e um comprimento médio dos caminhos mínimos menor do que o das redes k -Regulares equivalentes. A entropia pode ser ajustada por parâmetros do modelo de geração da rede. Portanto, as redes *small-world* podem ser classificadas entre redes regulares e redes aleatórias. Diversos modelos podem ser usados para gerar redes com o efeito *small-world*, mas um dos mais populares é o método de Watts-Strogatz (WS) [63]. Este método consiste em criar uma rede 2-regular e, a partir desta, criar $rp \cdot m$ enlaces, em que rp é uma probabilidade de religação. Esta criação de enlaces insere uma quantidade limitada de aleatoriedade na rede e permite a obtenção do efeito de atalhos na rede. Estes atalhos fornecem um meio de diminuir o diâmetro e o comprimento médio das redes 2-regulares para um valor mais próximo daquele observado em redes aleatórias, mas sem perder a característica de elevado coeficiente de agrupamento e baixa entropia das redes regulares. A distribuição do grau dos nós é semelhante a uma distribuição de Poisson. Estudos relacionaram este modelo a diversas redes do mundo real, como as redes sociais.

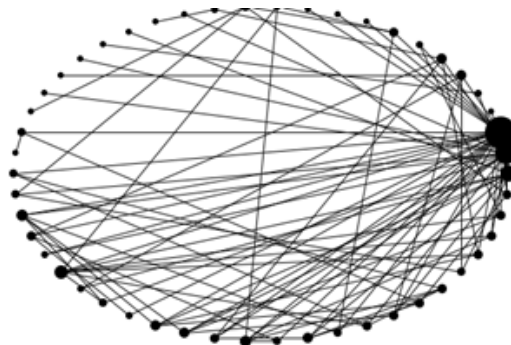
Uma rede livre de escala ou *scale-free* possui uma distribuição de grau dos nós que obedece



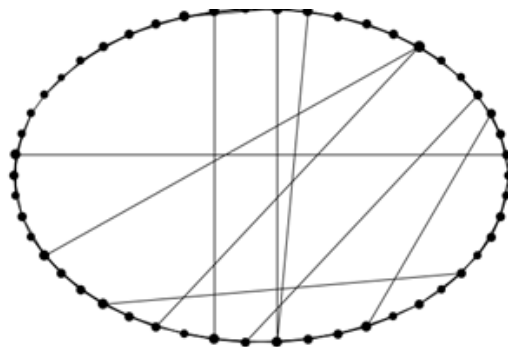
(a) Rede aleatória obtida pelo método ER.



(b) Rede 2-regular.



(c) Rede *scale-free* obtida pelo método BA.



(d) Rede *small-world* obtida pelo método WS.

Figura 3.2: Exemplos de redes criadas por quatro procedimentos de geração diferentes.

uma lei de potência na forma $h(d) \sim d^{-q}$, no qual d é o grau e q é um expoente tipicamente entre 2 e 3. Em outras palavras, uma rede que segue este modelo possui um pequeno número de nós com muitas conexões, denominados concentradores (*hubs*), mas a maioria dos nós da rede possui apenas uma ou poucas conexões. Redes *scale-free* possuem entropia elevada, mas abaixo das redes aleatórias equivalentes. Devido à presença dos *hubs*, as redes *scale-free* apresentam um valor de diâmetro e comprimento médio dos caminhos muito baixo. O método mais difundido para geração de redes *scale-free* é o método Barabási-Albert (BA) [63]. Neste método, o grafo é iniciado com três nós conectados e os nós seguintes são inseridos na rede usando o conceito de “ligação preferencial” (*preferential attachment*). A ligação preferencial significa que cada novo nó inserido na rede será ligado ao nó i com uma probabilidade

$$P(i) = \frac{d_i}{\sum_j d_j}, \quad (3.9)$$

em que d_i é o grau do nó i e $\sum_j d_j$ representa a soma do grau dos demais nós presentes na rede. Em outras palavras, os nós a serem ligados possuem uma maior tendência em se conectar aos nós com grau mais elevado. Logo, a assortatividade de redes *scale-free* é negativa. Estudos demonstraram que várias redes do mundo real, como a Internet, redes de estradas de ferro e sistemas biológicos, seguem este modelo [63].

Tabela 3.1: Comparação de propriedades topológicas de alguns modelos de rede para grafos com 50 nós, 100 arestas e $rp = 5\%$.

Propriedade	ER	BA	WS	2-regular
Assortatividade	-0,06	-0,24	0,07	1,00
CC	0,07	0,14	0,44	0,50
Distribuição	Poisson	Potência	Poisson	Delta(4)
Grau máximo	9	19	5	4
Grau médio	4,0	3,96	4,0	4
\bar{c}	2,83	2,68	4,54	6,63

A Tabela 3.1 fornece exemplos numéricos das medidas topológicas de redes com 50 nós e 100 enlases geradas por cada método explicado nesta seção. A Figura 3.2 apresenta um exemplo de rede aleatória (a), 2-regular (b), *scale-free* (c) e *small-world* (d) com as características das redes apresentadas na Tabela 3.1.

3.3 Proposta das Métricas FZC e HVC

3.3.1 Descrição

Embora os autovalores de matrizes características forneçam informações úteis, é difícil obter uma informação global de todos os autovalores por meio de uma métrica única.

Com o objetivo de melhor compreender o conjunto completo de autovalores da matriz Laplaciana de uma rede complexa, uma primeira proposta desta Tese consiste na aplicação da Transformada Discreta de Fourier - TDF (*Discrete Fourier Transform* - DFT) [73] sobre o conjunto completo de autovalores. A Figura 3.3(a) e a Figura 3.3(b) mostram as componentes pares e ímpares da DFT dos autovalores da matriz Laplaciana de uma rede livre de escala gerada pelo método BA [62]. Inicialmente, é possível observar que as componentes pares e ímpares da DFT apresentam um aspecto diferente quando observadas separadamente. É possível observar um padrão recorrente para outros modelos de redes complexas, como os modelos ER [60] e WS [61].

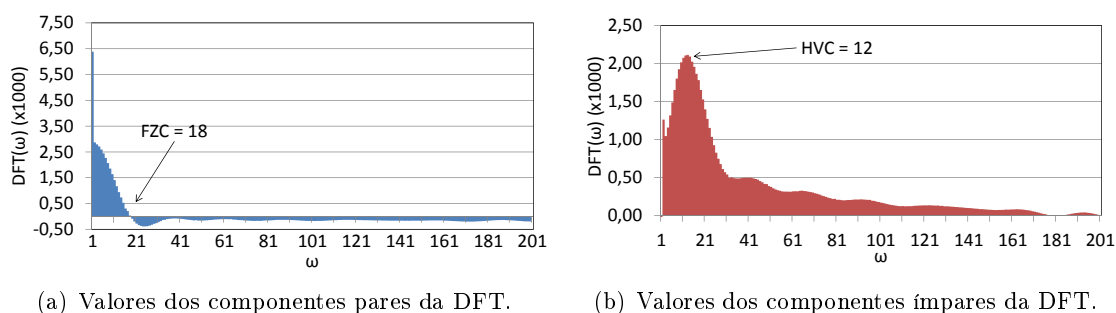


Figura 3.3: Valores dos componentes pares e ímpares da DFT dos autovalores da matriz Laplaciana de uma rede livre de escala.

Com base nas características observadas nas Figuras 3.3(a) e 3.3(b), foram propostas duas métricas simples envolvendo os valores das componentes da DFT. A primeira métrica, denominada FZC (*First Zero Component*), representa a frequência na qual a primeira componente par da DFT assume o valor zero, cruzando o eixo vertical. A segunda métrica, denominada HVC (*Highest Value Component*), representa a frequência relacionada com a componente ímpar de maior valor. O procedimento para calcular FZC e HVC está resumido no Algoritmo 1. No Algoritmo 1, a função $abs(\cdot)$ corresponde a $abs(z) = |z| = |-z| = z$. A função *senal* é definida por

$$senal(z) = \begin{cases} 1 & \text{para } z \geq 1 \\ -1 & \text{para } z \leq -1 \\ 0 & \text{para } z = 0 \end{cases} \quad (3.10)$$

O cálculo dos autovalores da matriz Laplaciana, ou seja, o passo 3 do Algoritmo 1, possui o maior custo computacional dentre os demais e, portanto, representa o gargalo no cálculo da métrica. O cálculo da matriz de grau e da Laplaciana possui complexidade $\mathcal{O}(n^2)$. Considerando que os cálculos dos autovalores é feito por meio do algoritmo QR [74], a sua complexidade computacional é $\mathcal{O}(n^3)$. A DFT é calculada usando o procedimento de Cooley-Tukey, e portanto, a complexidade computacional do passo 6 é $\mathcal{O}(n \log n)$ [73]. A separação dos valores

da DFT nos passos 5 e 6 pode ser feita usando uma repetição de complexidade $\mathcal{O}(n)$. Os laços 10-16 e 18-23 fornecem complexidade computacional $\mathcal{O}(n)$. Portanto, a complexidade computacional total do Algoritmo 2 é $\mathcal{O}(n^3 + 2n^2 + n \log n + 3n) = \mathcal{O}(n^3)$.

Algoritmo 1 Algoritmo para o cálculo de FZC e HVC.

```

1: procedimento CALCULARFZCHVC( $\mathcal{A}$ ,  $\mathcal{D}$ )
2:    $\mathcal{L} \leftarrow \mathcal{D} - \mathcal{A}$  ▷ Matriz Laplaciana
3:   Calcule os autovalores reais de  $\mathcal{L}$  e armazene em  $E_N$ 
4:   Calcule a DFT de  $E_N$  e armazene em  $DFT_N$ 
5:   Seja  $DFT_E$  o conjunto de componentes pares da DFT
6:   Seja  $DFT_O$  o conjunto de componentes ímpares da DFT
7:   Seja FZC a primeira componente de  $DFT_E$  que cruza o eixo em zero
8:   Seja HVC o componente com o maior valor de  $DFT_O$ 
9:   Seja lastValue =  $DFT_E(0)$ 
10:  para (int  $c = 0$ ;  $c < DFT_E.length$ ;  $c++$ ) faça
11:    se ( $abs(sinal(DFT_E(c)) + sinal(lastValue)) \neq 2$ ) então
12:      FZC  $\leftarrow c$ 
13:    pare
14:    fim se
15:    lastValue  $\leftarrow DFT_E(c)$ 
16:  fim para
17:  maxValue  $\leftarrow -\infty$ 
18:  para (int  $c = 0$ ;  $c < DFT_O.length$ ;  $c++$ ) faça
19:    se ( $DFT_O(c) > maxValue$ ) então
20:      maxValue  $\leftarrow DFT_O(c)$ 
21:      HVC  $\leftarrow c$ 
22:    fim se
23:  fim para
24: fim procedimento

```

3.3.2 Arranjo Experimental

Todos os resultados que foram obtidos nesta seção foram baseados em experimentos realizados em uma plataforma de simulação de redes complexas desenvolvida na linguagem de programação Java.

É desejável que as métricas calculadas sobre a DFT dos autovalores do Laplaciano apresentem valores diferentes para diferentes tipos de redes para que possam ser usadas de forma útil.

De modo a demonstrar a possibilidade de usar as métricas para classificação das redes, são geradas 45.000 redes diferentes usando os procedimentos de geração de redes de ER, BA e WS. O valor de densidade das redes (q) foi variado de 0,02 até 1,00 com um passo de 0,02. Para cada par (modelo de geração, q) são criadas 100 diferentes redes com tamanhos de 100, 200 e 400 nós.

As redes ER são geradas por meio da criação de enlaces entre pares (i, j) por meio de uma variável aleatória distribuída uniformemente com probabilidade p . O procedimento WS usado consiste na criação de um grafo k -regular e modificação dos enlaces existentes considerando uma probabilidade de religação de 0,10. O valor de q foi usado para obter o valor de k correspondente. Finalmente, as redes BA são geradas usando o procedimento de conexão preferencial. As redes BA são iniciadas com $n = 3$ e cada novo nó adicionado é conectado à rede pela adição de Δm novos enlaces. O valor de Δm também é obtido a partir de q . A probabilidade de um novo enlace ser conectado a um nó existente i pelo modelo BA é $P(i) \propto \frac{d_i^\tau}{\sum_{j=1}^n d_j^\tau}$, em que d_i é o grau do nó i e τ é uma constante para parametrização da conexão preferencial.

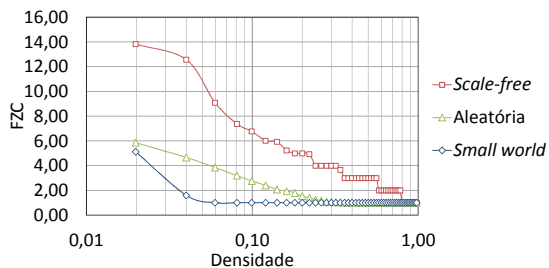
As métricas propostas também foram analisadas com relação à aplicação em redes reais de diferentes características. Nesta análise, redes aleatórias foram usadas a título de comparação (*benchmark*).

3.3.3 Resultados

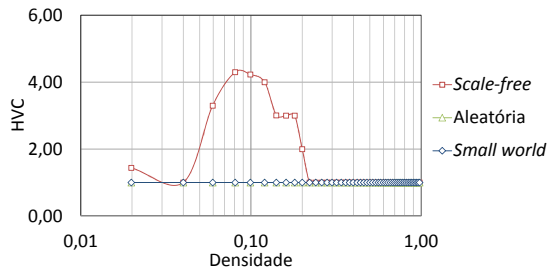
As Figuras 3.4 (a), 3.5 (a) e 3.6 (a) mostram o valor médio de FZC *versus* densidade, obtido a partir de 100 diferentes redes com 100, 200 e 400 nós, respectivamente. Todas as redes foram criadas de forma independente pelo respectivo modelo de geração. Ou seja, foi usado o modelo BA para criar redes *scale-free*, o modelo ER para criar redes aleatórias e o modelo WS para criar redes com efeito *small world*. Pode-se observar que FZC ocorre para altas frequências para redes com baixa densidade e diminui à medida que o valor de densidade aumenta, para todos os modelos. Além disso, redes livres de escala possuem o valor mais elevado de FZC quando são comparadas com redes de outros modelos e que possuem o mesmo valor de densidade.

As Figuras 3.4 (b), 3.5 (b) e 3.6 (b) mostram o valor médio de HVC *versus* densidade, obtido a partir de 100 redes com 100, 200 e 400 nós, respectivamente. As redes ER e WS possuem HVC= 1 para todas as redes geradas. Para as redes livres de escala, HVC= 1 ocorre apenas para redes com densidade acima de 0,20, independente do número de nós da rede. Pode-se ainda observar que o pico é deslocado à medida que o número de nós da rede aumenta.

Os resultados apresentados pelas Figuras 3.4 (a), 3.5 (a) e 3.6 (a) sugerem que é possível separar redes geradas por diferentes modelos canônicos, pois as redes geradas pelo modelo BA apresentam valores de FZC sempre acima dos valores obtidos para redes aleatórias e as redes geradas pelo modelo WS apresentam valores de FZC sempre abaixo dos valores obtidos para redes aleatórias. Ou seja, em princípio, é possível usar uma rede aleatória gerada pelo modelo ER para efeito de comparação (*benchmarking*), com o objetivo de identificar se uma dada rede real possui características mais próximas às de uma rede BA ou WS. Por outro lado, as Figuras 3.4 (b), 3.5 (b) e 3.6 (b) sugerem que o uso da métrica HVC é aplicável apenas para

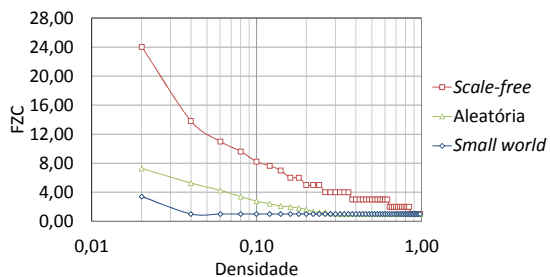


(a) Variação do valor de FZC.

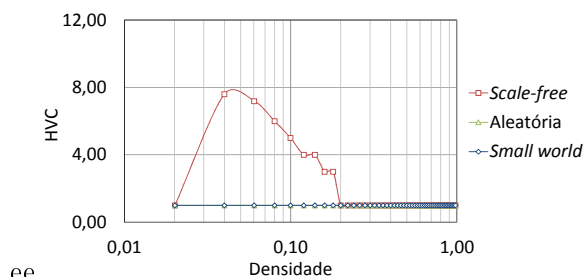


(b) Variação do valor de HVC.

Figura 3.4: Comparação de FZC (a) e HVC (b) de redes com 100 nós.

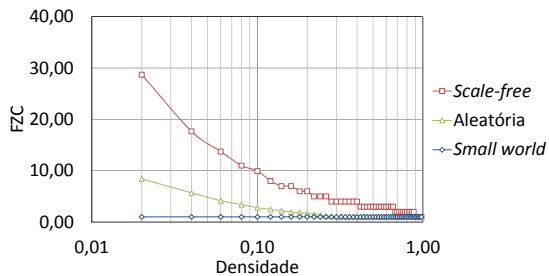


(a) Variação do valor de FZC.

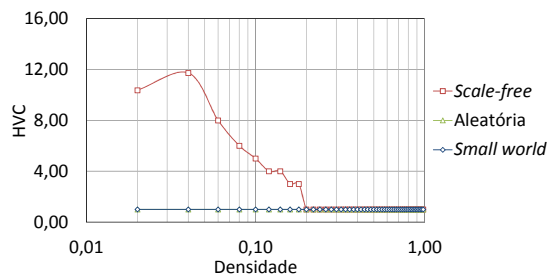


(b) Variação do valor de HVC.

Figura 3.5: Comparação de FZC (a) e HVC (b) de redes com 200 nós.



(a) Variação do valor de FZC.



(b) Variação do valor de HVC.

Figura 3.6: Comparação de FZC (a) e HVC (b) de redes com 400 nós.

diferenciação de uma rede puramente *scale-free* gerada pelo modelo BA das demais famílias de redes, bastando para tanto verificar se o valor de HVC é igual a 1,0. Além disso, este critério de classificação por HVC pode ser aplicado apenas a redes com $n > 100$ e para $q < 0,20$. Considerando que a métrica FZC se mostrou mais promissora para classificação das redes, o seu estudo será aprofundado.

Para uma aplicação efetiva da regra de decisão envolvendo FZC, é necessário aplicar um teste estatístico sobre um conjunto de amostras de redes de modo a permitir a validação da regra. Uma hipótese composta sobre a aplicação de FZC para classificação de redes de acordo com o modelo canônico de geração pode ser enunciada da seguinte forma: “Redes geradas pelo modelo ER possuem a mediana do valor de FZC acima dos valores obtidos para redes WS e abaixo dos valores obtidos para redes BA, para $0,02 < q < 0,20$ ”. Esta hipótese foi validada usando o teste não-paramétrico de Wilcoxon para pares de modelos canônicos com um nível de significância $\alpha = 0,05$ [75]. Para ilustrar a variação dos dados e a validação desta hipótese para $q = 0,02$, é apresentado um gráfico de *box-plot* para conjuntos de amostras com 100 redes na Figura 3.7. Na Figura 3.7, BA, WS e ER correspondem, respectivamente, aos modelos de geração para redes livres de escala, com efeito *small world* e aleatórias. *Min Outlier* e *Max Outlier* representam redes que apresentaram o valor da métrica abaixo de 1,5 vezes a altura da caixa e 1,5 vezes acima da altura da caixa, respectivamente. De acordo com a Figura 3.7, pode-se observar que a mediana da métrica FZC de redes BA está acima da mediana de FZC de todos os demais modelos de rede. Além disso, a Figura 3.7 também ilustra que a mediana de FZC de redes ER está acima da mediana de FZC referente ao modelo WS.

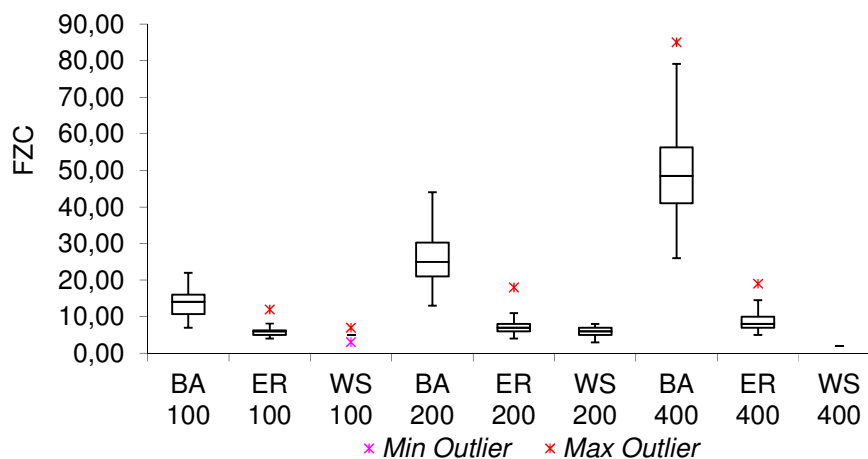


Figura 3.7: Gráfico de *box-plot* para análise da variação da métrica FZC para um conjunto de 100 amostras de cada tipo e para $q = 0,10$.

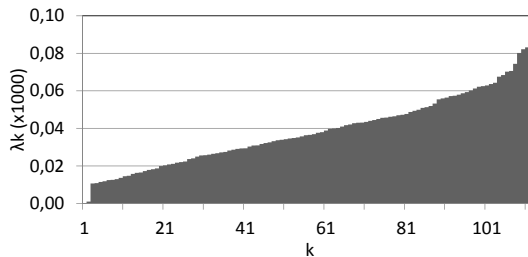
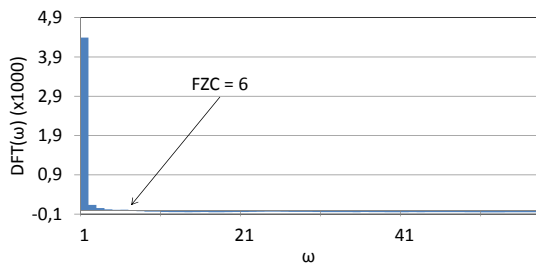
As métricas propostas também foram analisadas quando aplicadas a redes reais. A primeira rede real analisada é uma rede social que registra, por meio de ligações, o contato visual

entre pessoas durante uma conferência [76]. Esta rede apresenta características de redes livres de escala e o efeito de mundo pequeno. A segunda rede é a rede “Caenorhabditis elegans metabolic network” [77] (uma rede tipicamente livre de escala). Para enfatizar a expressividade da métrica FZC para caracterização de redes complexas, foram geradas redes usando os modelos BA, ER e WS com a mesmo tamanho e densidade das redes reais. Os valores de FZC e de outras métricas, tipicamente usadas para análise de redes complexas, são apresentados na Tabela 3.2. Com base nos dados da Tabela 3.2, é possível perceber que dentre as métricas apresentadas, FZC é a que mais distingue entre as características da rede original e as características das redes geradas pelo modelo canônico. Por exemplo, o uso de CC para classificação das redes não é apropriado pois CC obtido para a rede BA de 453 nós é igual à 0,042 e para a rede “Caenorhabditis elegans metabolic network”, $CC = 0,646$. Isso ocorre pois é possível gerar redes livres de escala com um valor ajustável de CC , mas em geral, redes BA tradicionais apresentam um coeficiente de agrupamento baixo. Por outro lado, o comprimento médio dos caminhos mínimos é muito semelhante quando são comparadas as redes de 113 nós geradas pelos modelos BA, ER, WS e a rede original “Sociopatterns Hypertext 2009”. Dentre as métricas tradicionais, o coeficiente de assortatividade é o que mais se aproxima de uma regra de classificação útil entre os três modelos canônicos. Contudo, dentre as quatro métricas expostas na Tabela 3.2, FZC é a que apresenta maior facilidade em classificar as redes originais como redes livres de escala, pois os valores de FZC obtidos para redes BA e para as redes originais são muito próximos.

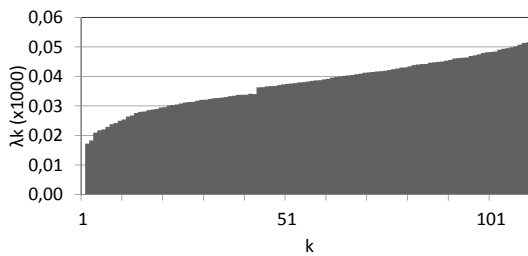
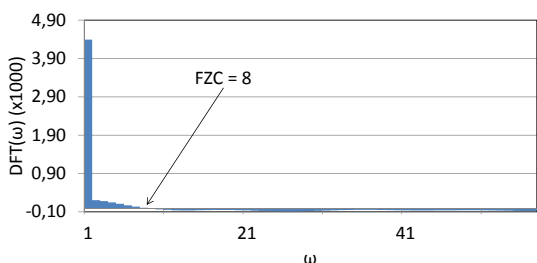
Tabela 3.2: Comparação entre redes reais e redes geradas por modelos canônicos que possuem o mesmo número de nós e densidade das redes originais.

Rede original analisada	Modelo	Assort.	CC	\bar{c}	FZC
Sociopatterns Hypertext 2009 (113 nós e $q = 0,35$)	BA	-0,132	0,425	1,650	8
	ER	-0,015	0,348	1,651	1
	WS	-0,087	0,601	1,698	1
	Original	-0,123	0,535	1,656	6
C. elegans metabolic (453 nós e $q = 0,02$)	BA	-0,092	0,042	2,880	57
	ER	-0,013	0,020	3,009	8
	WS	0,627	0,264	4,430	2
	Original	-0,220	0,646	2,664	56

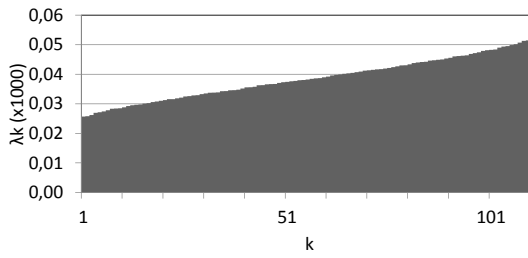
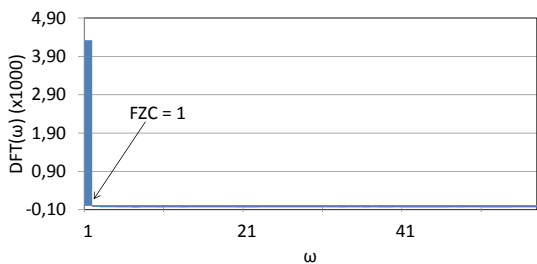
As Figuras 3.8 e 3.9 apresentam as componentes pares da DFT dos autovalores do Laplaciano e os autovalores em si, para as redes da Tabela 3.2. Em geral, os gráficos das componentes pares da DFT apresentam um comportamento semelhante, ou seja, possuem um pico nas primeiras frequências e em seguida o valor da DFT é rapidamente reduzido, muitas vezes apresentando uma alternância entre valores positivos e negativos ao longo do eixo das frequências. Contudo, uma inspeção visual das figuras sugere que os gráficos das componentes pares da DFT obtidos para as redes BA são mais semelhantes aos gráficos equivalentes das redes originais. Por exemplo, a Figura 3.8(c) é semelhante à Figura 3.8(a) pelo fato de apresentar um decaimento menos abrupto da DFT em comparação com as Figuras 3.8(e)



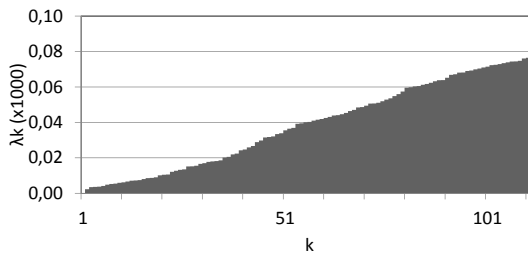
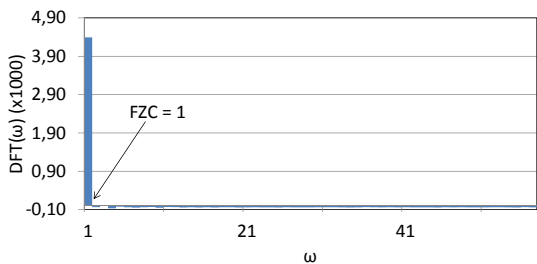
(a) Componentes pares da DFT para a rede original. (b) Autovalores da matriz Laplaciana para a rede original.



(c) Componentes pares da DFT para uma rede BA. (d) Autovalores da matriz Laplaciana para uma rede BA.

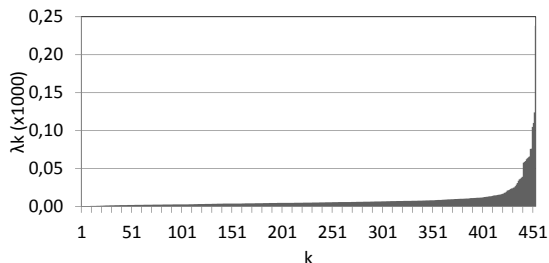
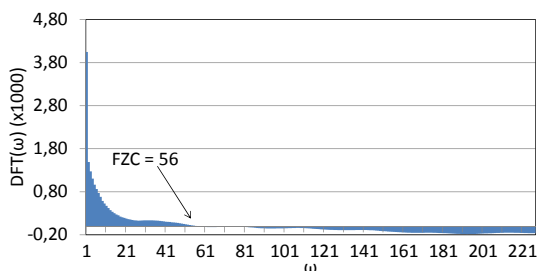


(e) Componentes pares da DFT para uma rede ER. (f) Autovalores da matriz Laplaciana para uma rede ER.

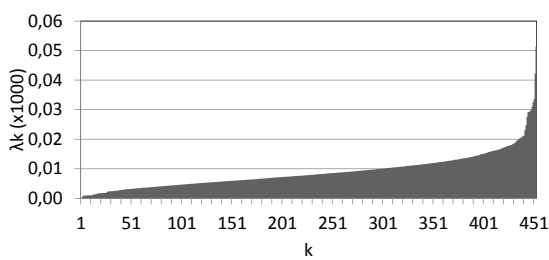
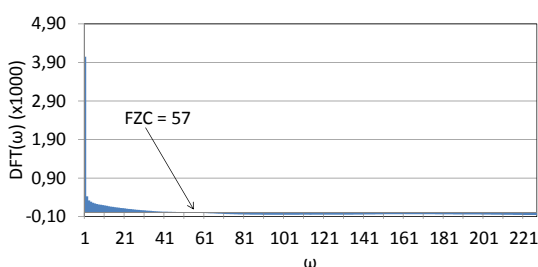


(g) Componentes pares da DFT para uma rede WS. (h) Autovalores da matriz Laplaciana para uma rede WS.

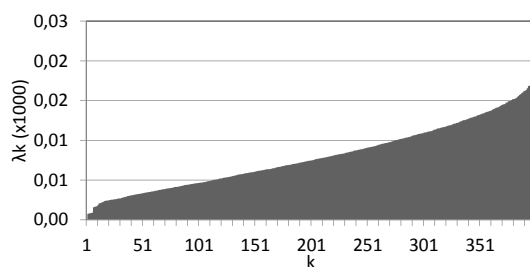
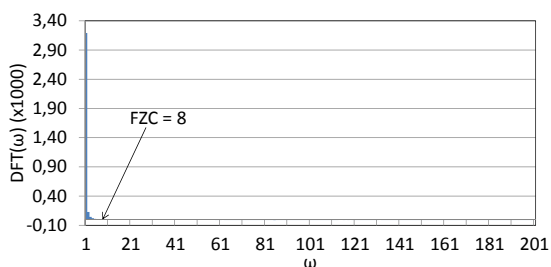
Figura 3.8: Autovalores da matriz Laplaciana e componentes pares da DFT dos autovalores para redes com mesmo número de nós de densidade da rede “Sociopatterns-Hypertext 2009”.



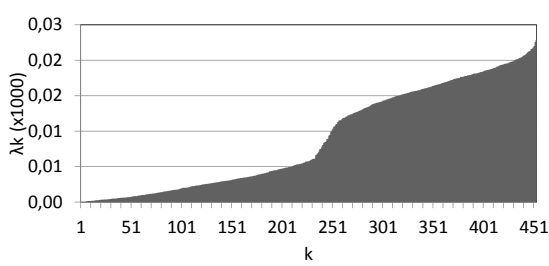
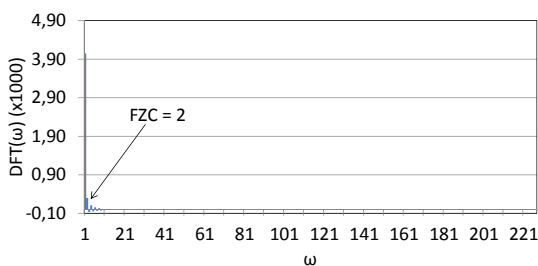
(a) Componentes pares da DFT para a rede original. (b) Autovalores da matriz Laplaciana para a rede original.



(c) Componentes pares da DFT para uma rede BA. (d) Autovalores da matriz Laplaciana para uma rede BA.



(e) Componentes pares da DFT para uma rede ER. (f) Autovalores da matriz Laplaciana para uma rede ER.



(g) Componentes pares da DFT para uma rede WS. (h) Autovalores da matriz Laplaciana para uma rede WS.

Figura 3.9: Autovalores da matriz Laplaciana e componentes pares da DFT dos autovalores para redes com mesmo número de nós de densidade da rede “Caenorhabditis elegans metabolic”.

e 3.8(g) e pelo fato de apresentar o primeiro cruzamento no eixo vertical próximo ao mesmo valor de frequência (definição de FZC). Por outro lado, os gráficos dos autovalores da matriz Laplaciana não permitem uma associação direta entre a rede original e os modelos canônicos. Por exemplo, apesar de a Figura 3.9(d) apresentar uma taxa de crescimento dos autovalores semelhante à apresentada na Figura 3.9(b), o mesmo não é possível de ser constatado quando são comparadas as Figuras 3.8(d) e 3.8(b). Além disso, o conjunto de autovalores não oferece um ponto único característico no qual poderia ser derivado uma métrica para classificação de redes de acordo com um modelo canônico de referência.

3.3.4 Considerações Finais sobre FZC e HVC

Os resultados apresentados nesta seção demonstram que o uso da Transformada Discreta de Fourier sobre os autovalores da matriz Laplaciana fornecem informações adicionais não retratadas por métricas tradicionais quando o objetivo é a classificação de redes com relação a um modelo canônico de referência.

Contudo, as métricas propostas nesta seção ainda apresentam algumas limitações que dificultam o seu uso prático. Em primeiro lugar, a métrica HVC apresenta o mesmo valor para redes WS e redes ER. Em segundo lugar, apesar de a métrica FZC se mostrar promissora para classificação das redes consideradas nos estudos de caso, a sua aplicação é restrita a uma faixa de densidade de enlaces que pode ser um entrave para sua aplicação no caso geral.

O estudo preliminar sobre o uso da TDF para avaliação de redes, incluindo a proposição de FZC e HVC, foi apresentado no “XXXI Simpósio Brasileiro de Telecomunicações”, em 2013 e publicado nos anais deste evento [78]. Apesar das limitações citadas nesta seção, este estudo preliminar inspirou o desenvolvimento de uma nova métrica, descrita na Seção 3.4.

3.4 Proposta da Métrica $I(\hat{\mathcal{F}})$

3.4.1 Descrição

A métrica proposta nesta seção é baseada na entropia dos coeficientes da Transformada de Fourier dos autovalores do Laplaciano (\mathcal{L}). \mathcal{L} contém um resumo útil da topologia da rede, pois contém informações sobre os graus dos nós e sobre os enlaces presentes na rede. Portanto, uma métrica derivada do espectro de \mathcal{L} pode resumir adequadamente a topologia da rede. Este aspecto já foi explorado de forma preliminar na Seção 3.3, mas as métricas topológicas propostas na Seção 3.3 e diversas outras encontradas na literatura não são válidas para determinados níveis de conectividade (densidade de enlaces). Além disso, uma métrica baseada na entropia contém informação sobre a aleatoriedade da rede, e contorna a limitação imposta para a métrica HVC (separação entre redes regulares e aleatórias).

Portanto, a proposta descrita nesta seção é usar a entropia dos coeficientes da DFT dos

autovalores do Laplaciano para classificar grafos de acordo com a topologia e o nível de aleatoriedade das redes. A métrica proposta pode ser calculada de acordo com o Algoritmo 2. A Equação (3.11) resume o cálculo numérico da métrica.

Algoritmo 2 O algoritmo usado para calcular $I(\hat{\mathcal{F}})$.

- 1: **procedimento** CALCULARIF
 - 2: Seja \mathcal{A} a matriz de adjacências do grafo G
 - 3: Calcule a matriz de grau \mathcal{D}
 - 4: Calcule a matriz Laplaciana $\mathcal{L} = \mathcal{D} - \mathcal{A}$
 - 5: Calcule os autovalores reais de \mathcal{L} e armazene em E
 - 6: Calcule a Transformada Discreta de Fourier (DFT) de E e armazene os valores em \mathcal{F}
 - 7: Normalize o conjunto \mathcal{F} de modo a obter valores entre 0 e 1 e armazene estes valores em $\hat{\mathcal{F}}$
 - 8: Calcule a entropia de $\hat{\mathcal{F}}$ de acordo com a Equação (3.11)
 - 9: **fim procedimento**
-

$$I(\hat{\mathcal{F}}) = - \sum_{i=1}^{|\hat{\mathcal{F}}|} (\hat{\mathcal{F}}_i \cdot \log_2 \hat{\mathcal{F}}_i). \quad (3.11)$$

A Eq. 3.11 possui uma estrutura semelhante à encontrada para o cálculo tradicional de entropia em teoria da informação e para a entropia dos graus dos nós, já conhecida em Ciência das Redes [63]. Ou seja, a métrica avalia a quantidade de informação, grau de incerteza ou previsibilidade sobre o valor numérico de uma determinada grandeza física. Contudo, a adaptação do conceito de entropia para o valor normalizado do conjunto \mathcal{F} remete simultaneamente à medida de incerteza sobre o grau dos nós e sobre a assortatividade de uma rede, visto que incorpora de forma simultânea: (i) o mecanismo de ligação entre os nós por meio da matriz Laplaciana (subsídio para avaliação de assortatividade); (ii) a avaliação de aleatoriedade sobre o grau dos nós por considerar que a entropia também atua sobre o grau dos nós que está presente na diagonal principal da matriz Laplaciana. Além disso, a mesma motivação apresentada aqui poderia ser aplicada diretamente ao conjunto dos autovalores, sem uso de uma transformação adicional (DFT). Contudo, os estudos realizados demonstram que a DFT proporciona uma análise coerente da métrica sobre uma faixa larga de densidades e a aplicação da entropia diretamente sobre o conjunto normalizado dos autovalores não possui uma mesma interpretação quando redes com densidades diferentes são consideradas.

A avaliação do Algoritmo 2 em termos de complexidade computacional é semelhante ao que foi apresentado para o Algoritmo 1. O cálculo dos autovalores da matriz Laplaciana também representa o gargalo no cálculo de $I(\hat{\mathcal{F}})$. Os passos 2 e 3 possuem complexidade $\mathcal{O}(n^2)$. Considerando que os cálculos dos autovalores é feito por meio do algoritmo QR [74], a sua complexidade computacional é $\mathcal{O}(n^3)$. A DFT é calculada usando o procedimento de Cooley-Tukey, e portanto, a complexidade computacional do passo 6 é $\mathcal{O}(n \log n)$ [73]. Os passos 7

e 8 fornecem complexidade computacional $\mathcal{O}(n)$. Portanto, a complexidade computacional total do Algoritmo 2 é $\mathcal{O}(n^3 + 2n^2 + n \log n + 2n) = \mathcal{O}(n^3)$.

3.4.2 Arranjo Experimental

Todos os resultados que foram obtidos nesta seção foram baseados em experimentos realizados em uma plataforma de simulação de redes complexas desenvolvida na linguagem de programação Java.

O objetivo das simulações é demonstrar, de forma experimental, que a entropia da DFT dos autovalores do Laplaciano apresenta diferentes valores para tipos diferentes de redes, considerando a densidade destas redes.

São geradas redes diferentes usando os procedimentos de geração de redes ER, BA e WS. São criadas redes com densidade q de 0,02 até 0,98 com um passo de 0,02. Para cada par (modelo de geração, q), são criadas 30 redes diferentes com os tamanhos de 100 e 1000 nós. As redes ER são geradas por meio da criação de enlaces entre pares (i, j) por meio de uma variável aleatória distribuída uniformemente com probabilidade p . O procedimento WS usado consiste na criação de um grafo k -regular e modificação dos enlaces existentes considerando uma probabilidade de religação de 0,10. O valor de q foi usado para obter o valor de k correspondente. Finalmente, as redes BA são geradas usando o procedimento de conexão preferencial. As redes BA são iniciadas com $n = 3$ e cada novo nó adicionado é conectado à rede pela adição de Δm novos enlaces. O valor de Δm também é obtido a partir de q . A probabilidade de um novo enlace ser conectado a um nó existente i pelo modelo BA é proporcional a: $P(i) \propto \frac{d_i^\tau}{\sum_{j=1}^n d_j^\tau}$, em que d_i é o grau do nó i e τ é uma constante de parametrização da conexão preferencial.

Para análise desta métrica também são usadas redes reais obtidas a partir de bases de dados de estudos anteriores em diversas áreas de aplicação (redes biológicas, redes de contato, redes de telecomunicações, redes de interação e redes sociais). Para cada rede real, foram geradas redes ER, BA e WS equivalentes (mantendo-se o número de nós e a densidade das redes originais). O valor de $I(\hat{\mathcal{F}})$ foi analisado com o objetivo de encontrar o melhor modelo para aproximação da rede real considerada. Depois deste passo, foram comparadas as propriedades topológicas das redes reais com as propriedades das redes geradas. Foram consideradas as seguintes bases de dados: Highland tribes [79]; Zachary karate club [80]; Hypertext 2009 [76]; Manufacturing emails [81]; Infectious [76]; Caenorhabditis elegans metabolic [77]; U. Rovira i Virgili [82].

Os resultados foram analisados com base em gráficos e tabelas dos valores obtidos para a métrica em diversos cenários. O comportamento estatístico dos valores da métrica proposta foi analisado por meio de gráficos de *box-plot* e por testes de hipótese considerando a mediana dos valores obtidos. Foi utilizado o teste não-paramétrico de Wilcoxon [75] para pares de modelos canônicos, com um nível de significância de 0,05.

3.4.3 Resultados

As Figuras 3.10(a) e 3.10(b) apresentam as curvas de “ $I(\hat{\mathcal{F}})$ versus densidade” para redes com 100 e 1000 nós, respectivamente. É possível observar que $I(\hat{\mathcal{F}}, d)$ apresenta o mesmo significado para redes pequenas e grandes. Redes BA apresentam os maiores valores de $I(\hat{\mathcal{F}}, d)$. Redes ER apresentam valores da métrica $I(\hat{\mathcal{F}}, d)$ abaixo dos valores das redes BA equivalentes e as redes WS apresentam os valores de $I(\hat{\mathcal{F}}, d)$ abaixo das curvas de ER e BA. $I(\hat{\mathcal{F}}, d)$ para redes k -Regulares assumem valores próximos de zero rapidamente. Este comportamento pode ser associado com a inexistência de aleatoriedade em redes k -Regulares e a capacidade desta métrica em capturar este tipo de informação. $I(\hat{\mathcal{F}}, d)$ de redes BA apresentam valores ainda maiores se o número de nós da rede é aumentado, mas para redes ER, o valor da métrica é pouco modificado quando o valor de n cresce, indicando que o modelo ER pode ser usado como *benchmark*. Se forem considerados dois valores distintos $I_A(\hat{\mathcal{F}}, d_1)$ e $I_B(\hat{\mathcal{F}}, d_1)$, relacionados com o mesmo valor de d mas diferentes modelos, A e B , observa-se para qualquer par de modelo que $I_A(\hat{\mathcal{F}}, d_1)/|I_A(\hat{\mathcal{F}}, d_1) - I_B(\hat{\mathcal{F}}, d_1)| > 0,10$ ($0 < d < 0,94$). Portanto, conclui-se que esta métrica é muito sensível às variações topológicas de cada modelo e ela pode ser usada adequadamente para classificar redes esparsas e densas e redes pequenas e grandes. Esta característica é uma vantagem sobre as métricas FZC e HVC, pois estas métricas são imprecisas para redes densas.

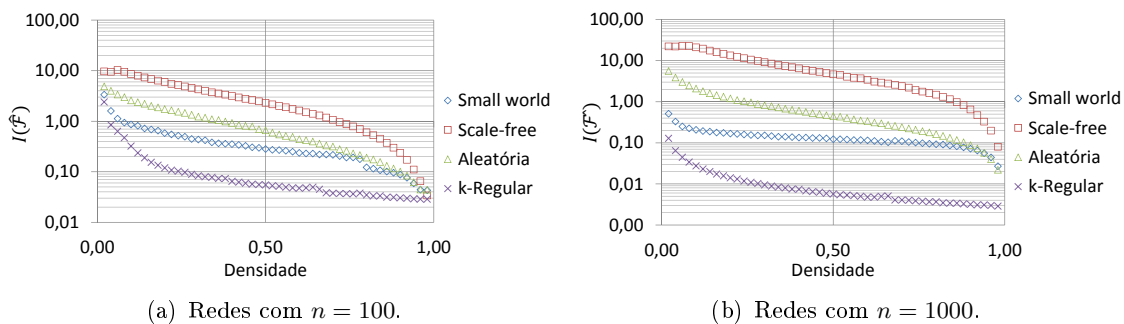


Figura 3.10: Métrica $I(\hat{\mathcal{F}})$ versus densidade para redes k -regulares, redes *scale-free* geradas pelo modelo BA, redes aleatórias geradas pelo modelo ER e redes *small-world* geradas pelo modelo WS.

Além disso, a Figura 3.11 ilustra alguns problemas encontrados quando outras métricas são usadas para avaliação de redes com diferentes densidades. A Figura 3.11(a) apresenta o gráfico de uma métrica conhecida na literatura, a entropia do grau dos nós ($I(G)$) [63], e explicita a variação desta métrica para diferentes densidades. É possível perceber que a avaliação do valor da métrica $I(G)$ comparativamente entre os modelos canônicos depende da faixa de densidade considerada. Por exemplo, se redes aleatórias forem usadas como *benchmarking* para avaliação de redes *scale-free*, esta avaliação deve considerar de forma diferenciada os intervalos $q < 0,10$, $0,10 < q < 0,96$ e $q > 0,96$. A Figura 3.11(b) apresenta o gráfico da

entropia dos autovalores de \mathcal{L} . Ou seja, a métrica da Figura 3.11(b) é semelhante à métrica proposta nesta seção, mas excetuando-se o uso da transformada de Fourier. Também é possível perceber com base na Figura 3.11(b) que não há uma forma prática de classificar as redes de acordo com o modelo canônico, pois a avaliação também depende da densidade considerada. Uma comparação visual entre as Figuras 3.11(a), 3.11(b) e a Figura 3.10(a) (proposta desta Tese) demonstra que apenas a última apresenta uma regra de classificação única, para uma ampla faixa de densidade.

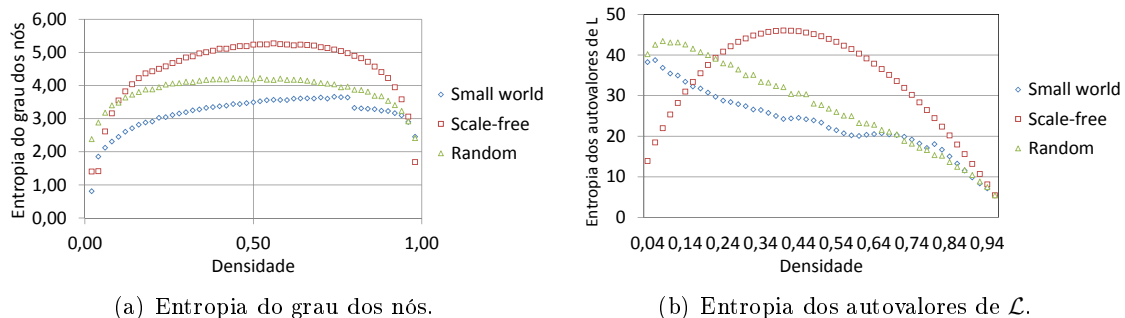


Figura 3.11: Entropia do grau dos nós e entropia dos autovalores de \mathcal{L} versus densidade para redes *scale-free*, redes aleatórias e redes *small-world*, de 100 nós.

Para aplicação prática da métrica para fins de classificação dos modelos canônicos, é importante avaliar a variação no cálculo da métrica quando redes semelhantes são apresentadas e aplicar um teste estatístico. Se for considerado um conjunto de redes com os mesmos valores de n , d e procedimento de geração, a métrica deve apresentar valores similares, pois as características topológicas são mantidas. Portanto, é possível enunciar a seguinte hipótese composta para validação da métrica $I(\hat{\mathcal{F}})$: “Redes BA apresentam valores de $I(\hat{\mathcal{F}})$ acima dos valores obtidos para redes ER e redes ER apresentam valores de $I(\hat{\mathcal{F}})$ acima dos valores obtidos para redes WS”. A Figura 3.12 apresenta gráficos de *box-plot* de $I(\hat{\mathcal{F}})$ para 30 diferentes redes com $n = 100$. A Figura 3.12(a) apresenta os dados para redes esparsas que foram geradas pelos algoritmos BA, ER e WS. É possível perceber que o valor mínimo para redes BA estão acima do valor máximo obtido para redes ER e o valor mínimo de redes ER está acima do valor máximo das redes WS. A Figura 3.12(b) resume a estatística descritiva para redes com $d = 0,30$. É possível perceber o mesmo comportamento observado para redes esparsas, mas a altura das caixas é menor. A Figura 3.12(c) apresenta a estatística descritiva para as redes densas. Conforme esperado, a variação no valor da métrica é reduzida para redes densas. Considerando que as sombras das caixas apresentadas na Figura 3.12 não são sobrepostas entre diferentes modelos, um teste estatístico é dispensável. Contudo, a aplicação do teste não-paramétrico de Wilcoxon para pares de modelos de redes, com nível de significância $\alpha = 0,05$, retifica que a hipótese enunciada para $I(\hat{\mathcal{F}})$ é válida.

A métrica $I(\hat{\mathcal{F}})$ também foi avaliada quanto à sua utilidade em avaliar redes reais dis-

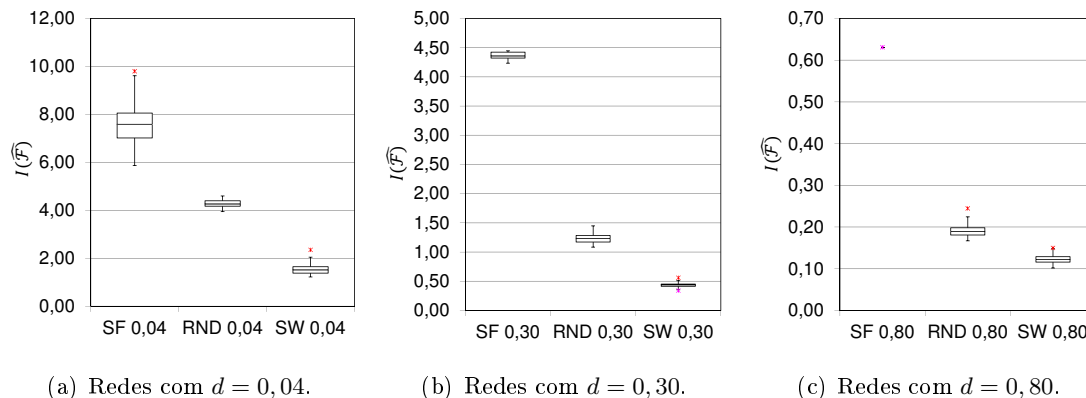


Figura 3.12: Box-plot para 30 redes diferentes com $n = 100$ e $d = 0,04$ (a); $d = 0,30$ (b) e $d = 0,80$ (c), respectivamente.

poníveis a partir de *datasets* de publicações anteriores. A Tabela 3.3 apresenta o valor de $I(\hat{\mathcal{F}})$ calculado para as redes reais e para cada modelo canônico. Na Tabela 3.3, foi considerado $rp = 0,05$ para redes WS e $\tau = 1$ para redes BA. Foi usado o valor de $I(\hat{\mathcal{F}})$ de cada modelo canônico para direcionar a escolha do modelo que melhor aproxima a rede real. Por exemplo, se $I(\hat{\mathcal{F}}) = 0,64$ para a rede “Highland tribes”, conclui-se que uma rede gerada pelo procedimento WS representa melhor esta rede do que o modelo canônico BA. Por outro lado, a rede “Infectious” é melhor representada por um modelo de rede livre de escala com uma “exponencial esticada”. Para efeito de comparação, foram criadas 30 redes com valores de $I(\hat{\mathcal{F}})$ entre o valor da rede original e o valor da rede gerada. É possível inferir importantes métricas topológicas com base neste procedimento. Por exemplo, a rede original “Highland tribes” possui $I(G) = 2,58$, $APL = 1,54$ e $k(hub) = 10$. Se forem calculados os valores médios de cada métrica para uma rede WS aproximada por $I(\hat{\mathcal{F}})$ ($rp = 0,15$), os valores obtidos são $I(G) = 2,33$, $APL = 1,54$ e $k(hub) = 10$. De acordo com estes resultados, é possível criar redes com $I(G)$, APL e grau do nó concentrador (*hub*) com um erro médio próximo a 0,10 para todas as redes analisadas se o procedimento de geração de redes que casa com o valor da métrica $I(\hat{\mathcal{F}})$ é usado.

3.4.4 Considerações Finais sobre $I(\hat{\mathcal{F}})$

A métrica $I(\hat{\mathcal{F}})$ foi proposta com o objetivo de fornecer um mecanismo prático para classificar redes em um dos modelos canônicos, de acordo com a semelhança em termos de propriedades topológicas. O cálculo da métrica envolve informações sobre o mecanismo de criação de enlaces na rede e sobre o grau dos nós. Estas informações são resumidas por meio da entropia do conjunto de valores da TDF dos autovalores de \mathcal{L} .

$I(\hat{\mathcal{F}})$ foi aplicada para classificação de redes com diferentes números de nós e diferentes densidades e foi demonstrado, por meio da análise da estatística descritiva dos experimentos,

Tabela 3.3: Redes reais e redes WS, ER e BA equivalentes.

Rede	n	d	$I(\hat{\mathcal{F}})$ para cada cenário				Gerador que melhor aproxima
			Original	WS	ER	BA	
Highland tribes	16	0,48	0,64	0,51	0,74	0,90	WS ($rp = 0, 15$)
Zachary karate club	34	0,13	2,76	0,90	1,82	2,67	BA ($\tau = 1, 29$)
Hypertext 2009	113	0,35	3,46	0,21	1,01	3,18	BA ($\tau = 1, 42$)
Manufacturing emails	167	0,23	5,60	0,37	1,64	5,43	BA ($\tau = 1, 07$)
Infectious	410	0,10	8,67	0,55	4,86	12,89	BA ($\tau = 0, 77$)
C. elegans metabolic	453	0,02	20,69	0,76	6,35	14,58	BA ($\tau = 1, 13$)
U. Rovira i Virgili	1133	0,01	18,19	0,71	8,55	26,09	BA ($\tau = 0, 79$)

que a métrica pode ser usada para classificar as redes geradas em um dos modelos estudados (BA, ER e WS). Além disso, a métrica proposta foi comparada com outras métricas disponíveis na literatura e aplicada para avaliação de diferentes *datasets* de redes publicados em estudos anteriores.

A proposição da métrica $I(\hat{\mathcal{F}})$ foi apresentada no “5th Workshop on Complex Networks”, em 2014. O trabalho foi publicado como um capítulo de livro da série “Studies in Computational Intelligence” [83]. $I(\hat{\mathcal{F}})$ foi usada em conjunto com outras três métricas topológicas para classificação de redes de transporte brasileiras e este estudo foi publicado na “Revista de Tecnologia da Informação e Comunicação” [84]. Além disso, foi realizado um estudo sobre o uso desta métrica para auxiliar na estimativa de probabilidade de bloqueio de redes ópticas. Este último estudo foi publicado no “Journal of Optical Communications and Networking” [85] e será detalhado no Capítulo 4.

3.5 Proposta de uma Métrica baseada em Concentração de Rotas (CR)

3.5.1 Descrição

Um grafo que representa uma rede usada para tráfego de informações entre os nós deve possuir uma topologia que facilite o fluxo de informações, considerando os recursos existentes na rede. Por exemplo, uma rede óptica com W comprimentos de onda apresenta níveis de desempenho diferentes se forem consideradas topologias distintas e cada uma possua uma capacidade diferente em distribuir o tráfego requerido. Se nesta rede óptica for usado o algoritmo de roteamento de menor caminho (*Shortest Path - SP*) e a topologia física exigir uma grande concentração de rotas passando por um mesmo enlace de fibra óptica, o desempenho da

rede será prejudicado devido ao bloqueio de chamadas pela ausência de canais disponíveis (comprimentos de onda). Desta forma, é desejável possuir uma métrica que possa quantificar a capacidade da topologia física de uma rede em distribuir a carga exigida para toda a rede.

Nesta seção será proposta uma nova métrica topológica baseada na avaliação de todas as rotas preferenciais para todos os pares (fonte, destino), que foi denominada de Concentração de Rotas (CR). O procedimento para calcular a métrica está detalhado no Algoritmo 3. Inicialmente, é definida uma matriz de concentração de rotas C com o tamanho $n \times n$. Em seguida, é necessário criar uma lista de rotas para cada par (fonte, destino). Se a métrica for usada para avaliação de um problema prático em rede, estas rotas precisam ser obtidas por meio do mesmo algoritmo de roteamento usado na rede estudada. O valor $C_{i,j}$ é atribuído com o número de caminhos ópticos que usam a fibra contida no enlace (i, j) . Finalmente, CR é definida por

$$CR = \frac{c_{max} - c_{min}}{r}, \quad (3.12)$$

em que c_{max} é o maior valor encontrado em C , c_{min} é o menor valor diferente de zero em C e r é o número de posições diferentes de zero em C . Embora não seja provável, é possível que alguns enlaces não sejam usados por nenhuma rota, então o algoritmo de cálculo da métrica estabelece que devem ser considerados apenas enlaces relacionados com posições diferentes de zero na matriz C . É possível observar que pequenos valores de CR sugerem redes com boa capacidade em distribuir tráfego. Por outro lado, valores elevados de CR correspondem a uma baixa capacidade em distribuir tráfego.

A relação da métrica com a capacidade da rede em distribuir tráfego pode ser facilmente compreendida se forem considerados pares de topologias físicas de redes com um mesmo número de nós e de enlaces, mas com valores diferentes para c_{max} e c_{min} . Em um primeiro cenário considere duas redes nas quais c_{min} possui o mesmo valor em ambas ($c_{min}^A = c_{min}^B$), mas em uma delas c_{max} é maior. Se $c_{max}^A > c_{max}^B$ significa que na rede A pelo menos um dos seus enlaces está mais sobrecarregado por rotas preferenciais de caminhos ópticos do que na rede B . O valor final da métrica para a rede A será comparativamente maior do que o valor da métrica para a rede B (considerando $c_{min}^A = c_{min}^B$). Neste primeiro cenário, é possível afirmar que a rede B possui uma maior capacidade em distribuir o tráfego da rede, pois $CR^B < CR^A$. Considere ainda um segundo cenário no qual duas redes possuem o mesmo valor para c_{max} , mas há uma diferença no valor de c_{min} . Se ($c_{min}^A < c_{min}^B$) significa que na rede A pelo menos um dos seus enlaces está mais livre com relação a rotas preferenciais de caminhos ópticos quando é comparada com a rede B . O valor final da métrica para a rede A será comparativamente maior do que o valor da métrica para a segunda rede (pois $c_{max}^A = c_{max}^B$). No segundo cenário, também é possível afirmar que a rede B possui uma maior capacidade em distribuir o tráfego da rede, pois $CR^B < CR^A$. Perceba que os cenários A e

B correspondem a gargalos de tráfego na topologia física devido a problemas distintos e em ambos a métrica foi capaz de capturar a melhor topologia física para fins de distribuição de carga. Outros cenários podem ser considerados com variações simultâneas em c_{max} e c_{min} , mas o resultado será semelhante aos casos considerados. A métrica oferecerá valores iguais de CR para topologias de rede distintas apenas nos casos em que c_{max} e c_{min} crescem ou diminuem na mesma proporção. Neste último cenário a métrica informa corretamente o fato de que as duas topologias possuem a mesma capacidade em distribuir tráfego. Perceba que o valor de r é usado para normalizar o valor numérico absoluto de $c_{max} - c_{min}$ e permitir a comparação da métrica inclusive para redes com características distintas.

Algoritmo 3 Algoritmo usado para calcular CR.

```

1: procedimento CR( $\mathcal{A}, n$ )
2:    $C \leftarrow \text{criarMatrizVazia}(n)$ 
3:   para  $fonte \leftarrow 0; fonte < n; fonte++$  faça
4:     para  $dest \leftarrow fonte + 1; dest < n; dest++$  faça
5:        $links \leftarrow \text{acharRota}(\mathcal{A}, fonte, dest)$ 
6:       para  $link \leftarrow 0; links < links.length; link++$  faça
7:         Incremente  $C[links[link].left][links[link].right]$ 
8:       fim para
9:     fim para
10:  fim para
11:   $c_{max} \leftarrow -\infty; c_{min} \leftarrow +\infty; r \leftarrow 0$ 
12:  para  $fonte \leftarrow 0; fonte < n; fonte++$  faça
13:    para  $dest \leftarrow source + 1; dest < n; dest++$  faça
14:      se  $C[fonte][dest] > 0$  então
15:        Incremente  $r$ 
16:        se  $C[fonte][dest] > c_{max}$  então
17:           $c_{max} \leftarrow C[fonte][dest]$ 
18:        fim se
19:        se  $C[source][dest] < c_{min}$  então
20:           $c_{min} \leftarrow C[fonte][dest]$ 
21:        fim se
22:      fim se
23:    fim para
24:  fim para
25:   $CR \leftarrow (c_{max} - c_{min})/r$ 
26: fim procedimento

```

A complexidade computacional do cálculo de CR depende do algoritmo de roteamento que é usado na rede óptica. Considere um algoritmo de roteamento por menor caminho que usa o algoritmo de Dijkstra. Se a implementação do algoritmo Dijkstra for baseada em fila de prioridades, a complexidade computacional obtida da literatura é $\mathcal{O}(e + n \log n)$. Considerando que o algoritmo de roteamento é acionado para todos os enlaces da rede, a

complexidade é multiplicada por um fator de $n(n-1)$ se as conexões são bidirecionais. Portanto, os passos 3-10 possuem complexidade computacional de $\mathcal{O}[n(n-1)(e+n \log n)]$. Os passos 12-24 possuem complexidade $\mathcal{O}(n^2)$. Portanto, a complexidade total do Algoritmo 3 é $\mathcal{O}\{[n(n-1)(e+n \log n)] + n^2\} = \mathcal{O}(n^4)$.

A métrica proposta nesta seção difere das que foram propostas nas Seções 3.3 e 3.4, principalmente devido à sua finalidade. As métricas FZC, HVC e $I(\hat{\mathcal{F}})$ foram propostas com o objetivo de resumir as propriedades topológicas das redes e permitir a sua classificação de acordo com um dos modelos canônicos estudados. A métrica CR, proposta nesta seção, é conhecida em Ciência das Redes como uma métrica de centralização de rede. Métricas de centralidade de nós e de arestas possuem a finalidade de avaliar a importância relativa destes elementos dentro da rede. Por exemplo, a centralidade de proximidade (*closeness*) mede a distância de um nó a todos os demais, considerando os caminhos mais curtos. Por outro lado, a centralidade de intermediação (*betweenness*) avalia o número de caminhos que passam por um determinado nó em particular. A centralização de uma rede na teoria das Redes Complexas é uma medida de quão central é o nó mais central da rede e fornece uma avaliação global de centralidade, considerando todos os nós da rede [1]. A métrica proposta nesta seção faz uso do conceito de centralidade de intermediação aplicada aos enlaces e aplica este conceito para o fornecimento de uma medida global para toda a rede que é mais simples do que a proposta de Freeman [1].

3.5.2 Arranjo Experimental

A métrica CR será avaliada de forma comparativa com outra métrica de centralização de rede disponível na literatura, para um conjunto de 1000 redes geradas a partir de cada modelo canônico. Para facilitar a análise neste estudo comparativo, foram consideradas redes de 50 nós e com densidade de enlaces que permita a geração de redes com grau médio 2 e 3, o que corresponde, para uma rede de 50 nós, aos valores de densidades $q = 0,04$ e $q = 0,06$, respectivamente. Este tipo de rede é comum para redes de telecomunicações devido aos requisitos de robustez. A análise comparativa é realizada por meio de gráficos de dispersão contendo os valores da métrica CR e da métrica de centralização por intermediação de Freeman descrita na Eq. 3.6.

De modo a compreender melhor a vantagem no uso da métrica CR, também são apresentados exemplos do cálculo da métrica e exemplos comparativos que apresentam redes com a mesma distribuição geográfica dos nós, mas com topologias diferentes. O objetivo do estudo é demonstrar que a métrica pode capturar informações de distribuições de tráfego que outras métricas não conseguem capturar. Esta avaliação é feita sobre a rede de 14 nós NSFNet exibida na Figura 3.13 considerando o algoritmo de roteamento por menor caminho.

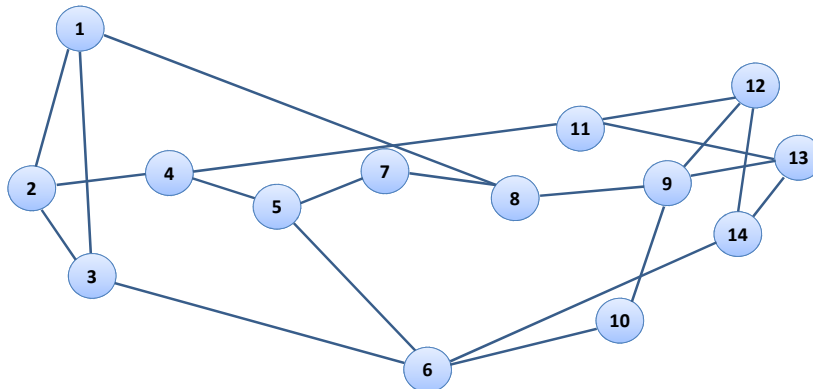


Figura 3.13: A rede NSFNet original, com 14 nós.

3.5.3 Resultados

A matriz C para a rede da Figura 3.13 está representada na Tabela 3.4.

Tabela 3.4: Matriz C para a rede NSFNet original.

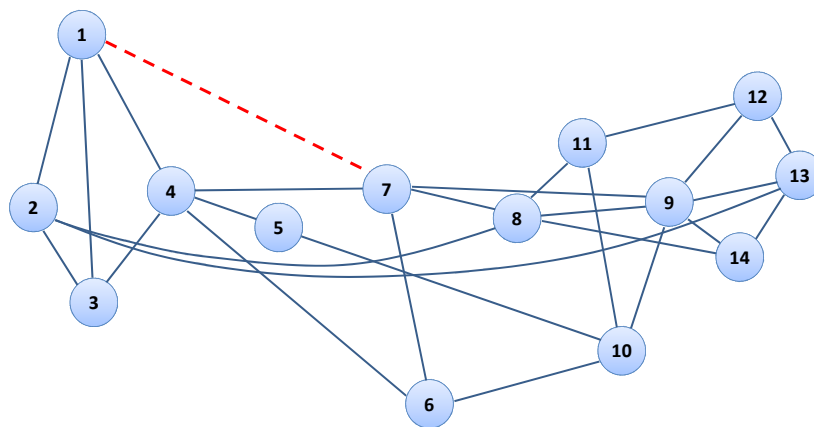
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	-	4	2	0	0	0	0	4	0	0	0	0	0	0
2	4	-	9	19	0	0	0	0	0	0	0	0	0	0
3	2	9	-	0	0	7	0	0	0	0	0	0	0	0
4	0	19	0	-	15	0	0	0	0	0	12	0	0	0
5	0	0	0	15	-	5	16	0	0	0	0	0	0	0
6	0	0	7	0	5	-	0	0	0	7	0	0	0	4
7	0	0	0	0	16	0	-	17	0	0	0	0	0	0
8	4	0	0	0	0	0	17	-	21	0	0	0	0	0
9	0	0	0	0	0	0	0	21	-	8	0	15	5	0
10	0	0	0	0	0	7	0	0	8	-	0	0	0	0
11	0	0	0	12	0	0	0	0	0	0	-	11	3	0
12	0	0	0	0	0	0	0	0	15	0	11	-	0	12
13	0	0	0	0	0	0	0	0	5	0	3	0	-	2
14	0	0	0	0	0	4	0	0	0	0	0	12	2	-

Cada valor $C_{i,j}$ na Tabela 3.4 fornece o número de demandas que usam o enlace (i, j) em seus caminhos ópticos (considerando o algoritmo de menor caminho por número de saltos). Por exemplo, $C_{2,4} = 19$ significa que o enlace 2-4 é usado por 19 demandas em seus caminhos ópticos. De acordo com a Tabela 3.4, o valor máximo é 21, o valor mínimo é 2 e existem 42 posições diferentes de zero ($C_{i,i} = 0, \forall i \in \{1, \dots, 14\}$). Portanto, $c_{max} = 21$, $c_{min} = 2$ e $r = 42$. Finalmente, é possível calcular CR para a rede NSFNet original (Figura 3.13) obtendo $CR = 0,45$.

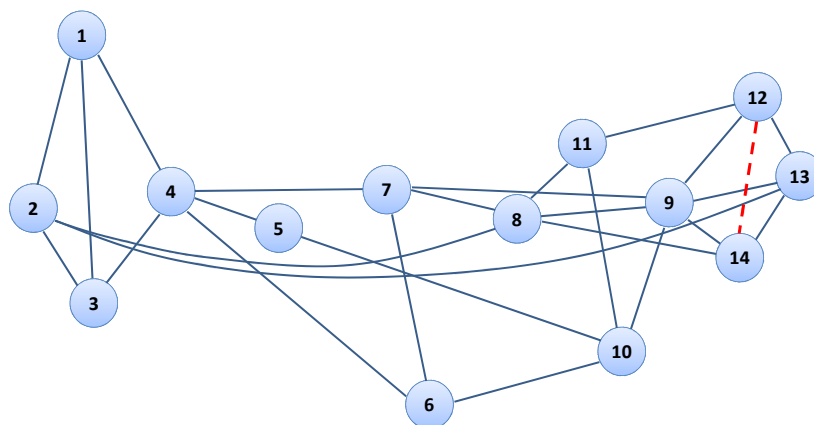
A Figura 3.14 exibe duas redes ópticas, a e b , sujeitas a uma carga de 200 erlangs. Estas redes são muito semelhantes, exceto pelos enlaces destacados. A rede da Figura 3.14(a) possui 40 comprimentos de onda por enlace, um fator de isolamento do ROADM de 38 dB, densidade $q = 0,31$, coeficiente de agrupamento $CC = 0,24$, comprimento médio dos caminhos $APL =$

83,37 km, $I(\hat{\mathcal{F}}) = 0,91$ e $CR = 0,42$. A rede da Figura 3.14(b) possui 40 comprimentos de onda por enlace, um fator de isolamento do ROADM de 38 dB, densidade $q = 0,31$, coeficiente de agrupamento $CC = 0,24$, comprimento médio dos caminhos $APL = 82,37$ km, $I(\hat{\mathcal{F}}) = 0,88$ e $CR = 0,68$. As redes foram avaliadas por um simulador de redes ópticas [11] e a ferramenta apresentou uma probabilidade de bloqueio $PB_a = 6,33 \cdot 10^{-3}$ e $PB_b = 2,41 \cdot 10^{-2}$, para as redes da Figura 3.14(a) e Figura 3.14(b), respectivamente.

Este exemplo serve para enfatizar que a proposição de novas métricas, como CR e $I(\hat{\mathcal{F}})$, permite uma avaliação mais precisa da rede, pois as outras cinco características apresentam valores iguais ou muito próximos. Além disso, o valor de probabilidade de bloqueio apresentado foi coerente com a capacidade de distribuição de tráfego sugerida por CR . Considerando que $CR(a) < CR(b)$, é esperado que a rede a possua um desempenho superior ao da rede b quando sujeita às mesmas condições. No Capítulo 4 serão fornecidos mais detalhes sobre o uso das métricas propostas neste capítulo para análise de redes WDM.



(a) Topologia física de rede WDM com $CR = 0,42$.



(b) Topologia física de rede WDM com $CR = 0,68$.

Figura 3.14: Comparação entre duas topologias semelhantes, mas que apresentam diferentes valores de CR .

A Figura 3.15 apresenta gráficos de dispersão para comparação da métrica proposta CR com a métrica de centralização de Freeman descrita na Eq. 3.6. As Figuras 3.15(a), 3.15(b) e 3.15(c) apresentam os valores das métricas para redes geradas pelos modelos ER, WS e BA, respectivamente. Cada ponto apresentado nestas figuras corresponde a uma diferente topologia de rede. As séries de dados $q = 0,04$ correspondem a redes com grau médio igual a 2 e as séries de dados $q = 0,06$ correspondem a redes com grau médio igual a 3. Uma análise inicial nas três figuras sugere que as duas métricas apresentam correlação positiva, ou seja, há uma tendência na qual quando o valor de centralidade de Freeman cresce, o valor de CR também cresce. A correlação positiva é desejável, visto que as duas métricas propõem avaliar uma característica semelhante nas redes. O cálculo do coeficiente de correlação de Pearson [75] para as redes que possuem $q = 0,04$ fornece $r = 0,91$, $r = 0,99$ e $r = 0,99$ para os modelos ER, WS e BA, respectivamente. O cálculo do coeficiente de correlação de Pearson para as redes que possuem $q = 0,06$ fornece $r = 0,91$, $r = 0,97$ e $r = 0,94$ para os modelos ER, WS e BA, respectivamente.

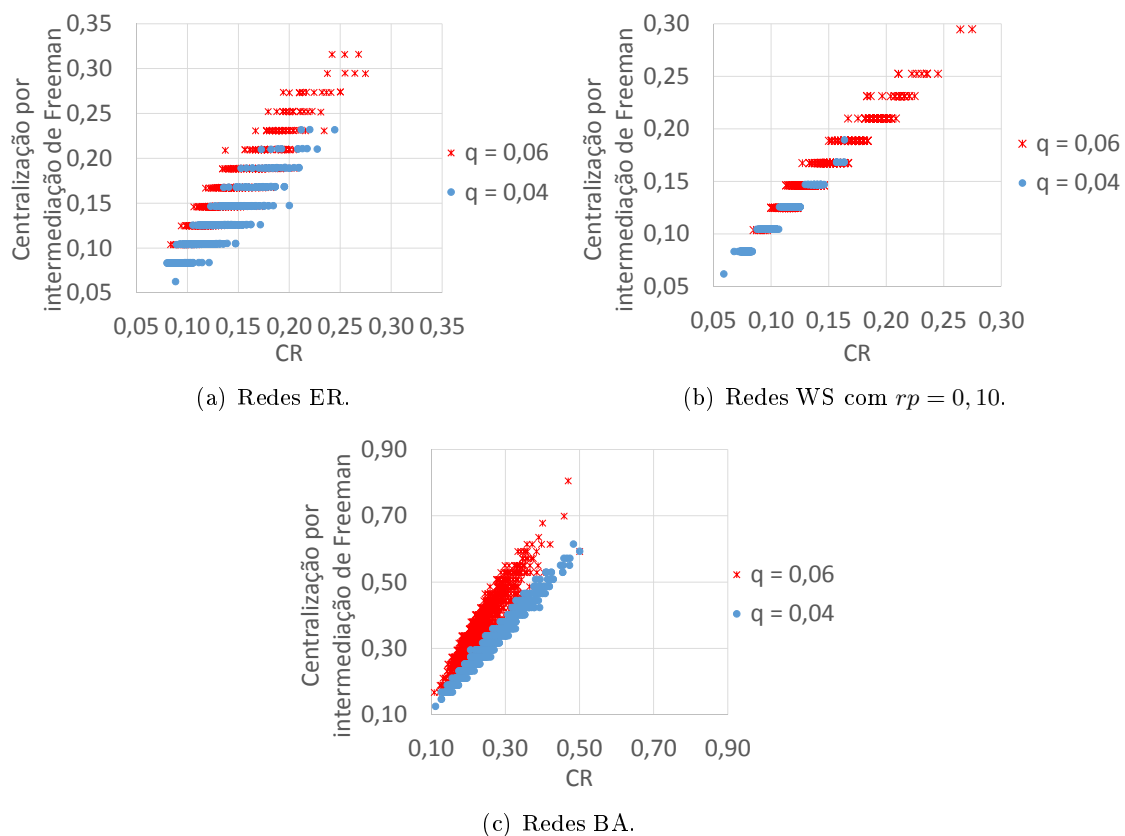


Figura 3.15: Gráficos de dispersão para avaliação da correlação entre a métrica proposta CR e a métrica de centralização de Freeman [1].

Além disso, é possível perceber, com base na Figura 3.15, que ambas as métricas conseguem

avaliar a capacidade da rede em facilitar o fluxo de informações. Por exemplo, redes geradas pelo modelo BA possuem a tendência de gerar concentradores (*hubs*) e este comportamento promove uma concentração no fluxo de informações. Uma comparação entre a Figura 3.15(c) e as Figuras 3.15(a) e 3.15(b) revela que as redes BA de fato apresentam valores mais elevados das duas métricas em comparação com as redes ER e WS. Além disso, como as redes WS são mais regulares do que as demais, o valor de concentração de caminhos foi menor para este modelo, dentre os três modelos analisados. Também é possível perceber, para os três modelos, que no conjunto de redes com densidade $q = 0,06$ (grau médio 3) há redes com concentração de caminhos maiores do que as redes obtidas para $q = 0,04$ (grau médio 2).

Uma última constatação com relação à comparação da métrica CR com a métrica de Freeman é que esta última apresenta o mesmo valor de centralização para um conjunto de redes diferentes, enquanto que a nova métrica consegue distinguir algumas destas redes. Por exemplo, ao analisar a Figura 3.15(a), é possível perceber diversos pontos com o mesmo valor no eixo vertical e valores distintos no eixo horizontal. Esta característica pode ser explicada pelo fato de a métrica de Freeman ser calculada com base em uma soma de diferenças de centralidade dos nós/arestas do grafo e pelo fato de haver diversas combinações de centralidade de nós/arestas que oferecem o mesmo valor final da métrica. Por exemplo, considere uma rede A na qual o valor máximo de centralidade por intermediação é 10 e uma rede B na qual o valor máximo de centralidade por intermediação é 8. Há diversas combinações de centralidade dos demais nós/arestas das redes A e B que podem levar o valor de centralização destas redes a ser idêntico. De fato, considere a situação hipotética na qual todos os demais elementos da rede A possuem centralidade igual a 6 e todos os demais elementos da rede B possuem centralidade igual a 4. Neste cenário hipotético estas duas redes passam a ter o mesmo valor de centralização, mesmo considerando que a rede A possui um valor mais elevado de centralidade máxima.

3.5.4 Considerações Finais sobre CR

A métrica CR foi proposta nesta seção como uma alternativa prática para avaliação da facilidade que uma dada topologia de rede possui em distribuir um fluxo de informações, considerando uma métrica de centralidade baseada em intermediação. Foi mostrado, por meio de exemplos de topologias, que a métrica consegue diferenciar um grafo que possui dificuldade em distribuir o tráfego de um grafo que apresenta maior facilidade para distribuição de tráfego.

A métrica proposta também foi comparada com uma métrica conhecida da literatura e que possui finalidade semelhante. Foram usados gráficos de dispersão para demonstrar que ambas as métricas podem ser usadas para avaliação da concentração de caminhos em uma rede. Contudo, a comparação entre as métricas sugere que CR consegue distinguir alguns conjuntos de redes que apresentam um mesmo valor de centralização pela definição de Freeman.

A proposição da métrica CR foi publicada no “*Journal of Optical Communications and*

Networking” [85]. Além disso, nesta mesma publicação foi apresentado um estudo sobre o uso desta métrica para auxiliar na estimativa de probabilidade de bloqueio de redes ópticas. Este último estudo será detalhado no Capítulo 4.

Capítulo 4

Análise de Redes Ópticas

Conforme apresentado no Capítulo 2, redes ópticas com multiplexação de comprimentos de onda usam conexões comutadas por circuito conhecidas como caminhos ópticos (*light-paths*). Estes caminhos ópticos são roteados por meio de dispositivos de núcleo, tais como os *Optical Add/drop Multiplexers* (OADMs) ou *Wavelength Selective Switches* (WSS) [2]. Para o estabelecimento de um novo caminho óptico, um algoritmo de Roteamento e Atribuição de Comprimento de Onda (*Routing and Wavelength Assignment algorithm* - RWA) precisa ser usado para selecionar uma rota e um comprimento de onda (λ) livre para atender uma requisição de chamada. Em redes que possuem tráfego dinâmico, algumas das requisições de chamadas podem não ser atendidas. Isto pode ocorrer devido à indisponibilidade de um caminho óptico, seja devido à falta de comprimentos de onda ou aos critérios de Qualidade de Transmissão (*Quality of Transmission* - QoT) pré-estabelecidos. Em geral, a probabilidade de bloqueio (PB) é usada para avaliação do desempenho de uma rede óptica sujeita a tráfego dinâmico.

PB pode ser estimada por expressões analíticas [34, 35, 36] ou por simuladores de eventos discretos [11]. Em geral, é muito difícil determinar o valor exato de PB para uma rede usando uma expressão matemática fechada, pois a PB está intimamente relacionada com a dinâmica da rede, incluindo a distribuição de tráfego e o número de comprimentos de onda usados em cada enlace. Além disso, métodos baseados em expressões analíticas geralmente ignoram as penalidades da camada física, considerando um número fixo de comprimentos de onda em cada enlace ou assumindo restrições rígidas para a topologia física da rede. Em geral, estes métodos podem lidar adequadamente apenas com redes lineares ou em anel e em cenários muito simples.

Por outro lado, simuladores de eventos discretos são ferramentas poderosas para a estimativa de PB. Simuladores fornecem uma alternativa prática para incluir a avaliação das penalidades da camada óptica na estimativa do desempenho da rede. Apesar das vantagens, os simuladores de eventos discretos frequentemente apresentam um elevado custo computaci-

onal devido ao grande número de chamadas que é necessário para estimar PB com precisão. Uma alternativa diferente foi proposta em [86], denominada *Interactive Matrix Methodology* (IMM). Embora IMM seja um método rápido, na proposta [86] também não são consideradas as penalidades da camada física.

Conforme apresentado no Capítulo 3, diversos estudos sugerem que é possível aplicar as teorias desenvolvidas em Ciência das Redes para análise de redes reais. Uma das hipóteses levantadas nesta Tese é a de que é possível mapear as relações complexas existentes entre as características de uma rede óptica (incluindo as propriedades topológicas) e o seu desempenho em termos de PB. Este mapeamento poderá viabilizar ferramentas de análise precisas e de baixo custo computacional. Uma ferramenta com estas características permitiria a análise de grandes redes de transporte em malha, considerando penalidades da camada física. Este capítulo propõe o uso de um analisador de redes ópticas que usa em conjunto os conceitos de Ciência das Redes e técnicas de aprendizagem de máquinas para mapear as propriedades da rede em um valor de PB. Esta nova abordagem fornece uma ferramenta precisa e rápida para a aferição do desempenho de redes ópticas.

O Capítulo está organizado da seguinte forma. A Seção 4.1 descreve o funcionamento de um simulador de eventos discretos para análise de redes, o SIMTON [11]. O objetivo desta seção é fornecer ao leitor um melhor entendimento sobre o funcionamento de simulador de eventos discretos típico e facilitar o entendimento da proposta desta tese para análise de redes. A Seção 4.2 apresenta a nova proposta deste trabalho para análise de redes ópticas, incluindo uma descrição detalhada da proposta, o arranjo experimental usado para avaliação e os resultados obtidos. A Seção 4.2.4 apresenta considerações finais sobre as contribuições da Tese em análise de redes e resume as publicações relacionadas.

4.1 Simulador de Eventos Discretos

Neste trabalho foi usado um simulador de eventos discretos pré-existente, denominado SIMTON (*Simulator for Transparent Optical Networks*) [11]. O SIMTON é um simulador de redes ópticas originalmente implementado na linguagem de programação C++. Ele é constituído de uma interface gráfica e um motor de simulação. O motor de simulação também possui uma versão na linguagem de programação JAVA, a qual foi desenvolvida a partir da versão original em C++, contemplando algumas otimizações, visando melhorar o desempenho em termos de tempo de execução, para obtenção dos resultados contidos em [87]. O motor de simulação considera o modelo de camada física proposto por Pereira e colaboradores em [30]. Nesta tese é usado o motor de simulação do SIMTON desenvolvido na linguagem JAVA [87]. Nas próximas subseções serão detalhados os aspectos mais importantes do SIMTON.

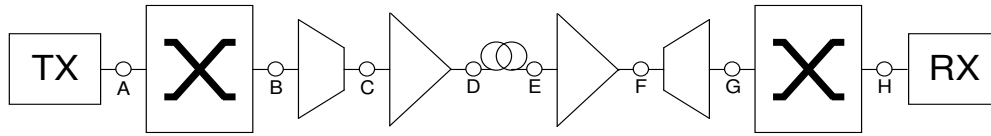


Figura 4.1: Estrutura de um enlace em uma rede WDM considerada pelo SIMTON.

4.1.1 Modelo de Camada Física

O SIMTON permite a simulação de redes ópticas WDM. Conforme mencionado no Capítulo 2, uma rede WDM consiste em nós conectados por fibra óptica e cada enlace possui diversos comprimentos de onda que podem ser utilizados para transportar dados. Se usuários de dois nós precisam transportar dados, é feito um pedido de conexão à rede óptica. O pedido é analisado pelo plano de controle da rede, denominado de Controle de Admissão de Chamadas (CAC). Se o pedido for aceito, é estabelecido um caminho físico denominado caminho óptico (*lightpath*). Portanto, assumem-se aqui redes ópticas comutadas por circuito [88]. É importante destacar que o SIMTON, ao simular uma rede óptica transparente, usa a premissa de que uma rota composta por um caminho óptico entre dois pontos A e B usa o mesmo comprimento de onda. Na literatura, essa restrição é conhecida como restrição de continuidade de comprimento de onda (RCCO) [2].

O SIMTON também assume enlaces bidirecionais. Ou seja, se existe uma ligação física entre dois nós A e B, há uma fibra óptica para o sentido ascendente e outra fibra para o sentido descendente. Um enlace da rede óptica possui uma estrutura predeterminada no SIMTON e essa estrutura é exibida na Figura 4.1. Da esquerda para a direita os dispositivos são: lasers transmissores, comutador óptico, multiplexador, amplificador óptico, fibra óptica, amplificador óptico, demultiplexador, comutador óptico e receptores. Os ganhos dos amplificadores ópticos são calculados para compensar as perdas nos enlaces e nos nós. Para a simulação de uma rede transparente, o SIMTON considera a arquitetura de nó mostrada na Figura 4.2.

A formulação utilizada pelo SIMTON para avaliação das penalidades da camada física é uma combinação da formulação proposta por Pereira *et al.* em [30] para avaliação de OSNR com as propostas apresentadas em [89, 90] para avaliação da dispersão do pulso óptico. As penalidades consideradas pelo SIMTON são as seguintes: perdas nos multiplexadores, nos demultiplexadores e nos comutadores ópticos; ruído de emissão espontânea da fonte (SSE - *Spontaneous Source Emission*) no transmissor óptico; mistura de quatro ondas (FWM), dispersão por modo de polarização (PMD) e dispersão cromática residual (RD - *Residual Dispersion*) na fibra óptica de transmissão; ruído (ASE) no amplificador EDFA; a saturação de ganho no amplificador EDFA e o *crosstalk* no comutador óptico. Para cada caminho óptico LP_i , a implementação do modelo de camada física dentro do SIMTON retorna dois valores:

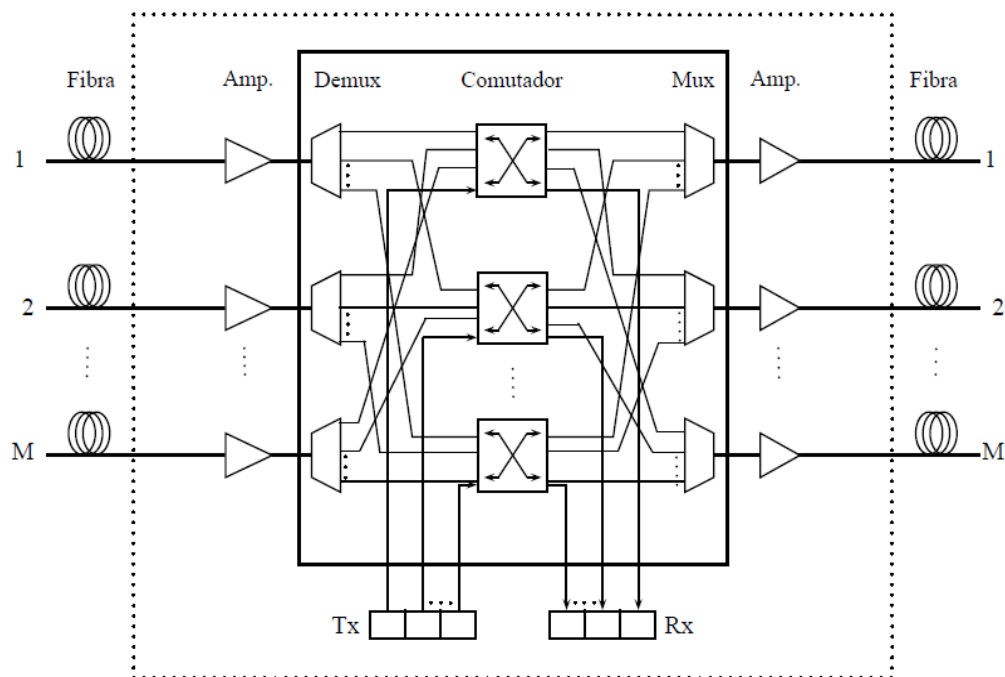


Figura 4.2: Arquitetura de um nó transparente para as simulações do SIMTON.

$OSNR(LP_i)$, que representa a OSNR no fim do caminho óptico LP_i (no receptor) e $\Delta_t(LP_i)$, que representa o alargamento temporal percentual do pulso óptico no final do caminho óptico LP_i (no receptor). Quanto maior for $OSNR(LP_i)$ e menor for $\Delta_t(LP_i)$, menos penalizado está o sinal que se propaga no caminho óptico LP_i . A descrição da formulação quantitativa das penalidades feitas pelo SIMTON não será detalhada aqui. Uma explicação detalhada sobre as equações usadas pelo SIMTON para o cálculo de cada penalidade pode ser obtida em [21].

4.1.2 Geração de Chamadas

O SIMTON assume que os pedidos de conexão à rede chegam de forma dinâmica DLE (DLE - *Dynamic Lightpath Establishment*) [91], um seguido do outro. O processo de geração dos pedidos de conexão é um processo estocástico. Para simular os diversos pedidos de conexão requisitados à rede, o SIMTON possui um gerador de chamadas que gera uma quantidade de pedidos de conexões previamente configurada. Para cada chamada, dois nós da rede são escolhidos por meio de uma distribuição uniforme, para simularem pares de origem e destino que desejam de comunicar.

O processo de chegada de pedidos de conexão segue um processo de Poisson: o intervalo de tempo entre chamadas é exponencialmente distribuído com média $1/\mu$ e a duração de cada chamada também segue uma distribuição exponencial, porém com média H . Define-se $1/\mu_{i,j}$

como o intervalo de tempo médio entre chamadas com origem no nó i e destino no nó j e $H_{i,j}$ como a duração média das chamadas que se originam no nó i e terminam no nó j . A carga $L_{i,j}$ das conexões entre i e j é definida por $L_{i,j} = \mu_{i,j} \cdot H_{i,j}$ [21]. O SIMTON pressupõe um tráfego simétrico, ou seja, $L_{i,j} = L_{j,i}$. A carga total da rede L_{rede} é definida por

$$L_{rede} = 2 \sum_{i=1}^n \sum_{j=i+1}^n \mu_{i,j} \cdot H_{i,j}, \quad (4.1)$$

em que n é o número total de nós na rede. O perfil de tráfego da rede depende dos parâmetros $\mu_{i,j}$ e $H_{i,j}$. Se $\mu_{i,j} = \mu \forall i, j$ e $H_{i,j} = H \forall i, j$ o tráfego é definido como **uniforme**. Caso contrário, o tráfego é definido como **não-uniforme**. Apesar de L_{rede} ser adimensional, na literatura é comum se atribuir a unidade erlang à carga da rede. A carga da rede é uma medida de quanto a rede está sendo utilizada: alta carga significa que existem muitas conexões em curso ao longo de rede, enquanto que baixa carga significa que existe um número pequeno de conexões ativas na rede.

4.1.3 Controle de Admissão de Chamadas (CAC) e Estimativa de PB

O controle de admissão de chamadas é parte do plano de controle da rede óptica e é quem decide se um pedido de conexão pode ser atendido ou não. O CAC do SIMTON é baseado em dois aspectos: se há comprimento de onda disponível para o caminho óptico e se o caminho óptico encontrado para implementar o pedido de conexão atende aos requisitos mínimos de qualidade de transmissão (QoT) pré-estabelecidos.

A Figura 4.3 ilustra a avaliação de uma requisição de chamada à rede óptica. Dado um par (origem, destino), uma rota é selecionada utilizando um algoritmo de roteamento. Nesta tese utilizou-se apenas o algoritmo de menor caminho por distância calculado pelo algoritmo de Dijkstra [88]. Após a seleção da rota, o motor de simulação tenta obter um comprimento de onda usando um algoritmo de alocação de comprimento de onda. Nesta Tese, utilizou-se apenas o algoritmo *first fit*, que seleciona o primeiro comprimento de onda livre de uma lista pré-definida. A relação sinal ruído óptica (OSNR) é então calculada com base no caminho óptico selecionado. Se a OSNR estiver acima de um nível predeterminado ($OSNR_{QoT}$), a chamada é estabelecida, caso contrário a chamada é bloqueada. A chamada também é bloqueada caso o alargamento do pulso óptico (Δt) esteja acima de um nível máximo (δ) ou se não houver comprimento de onda disponível.

A probabilidade de bloqueio é uma medida da quantidade relativa de pedidos de conexão que não são aceitos pelo CAC. Ou seja, quanto maior a probabilidade de bloqueio, menos usuários são atendidos e, conseqüentemente, pior é o desempenho desta rede. PB é estimada pelo SIMTON por meio de

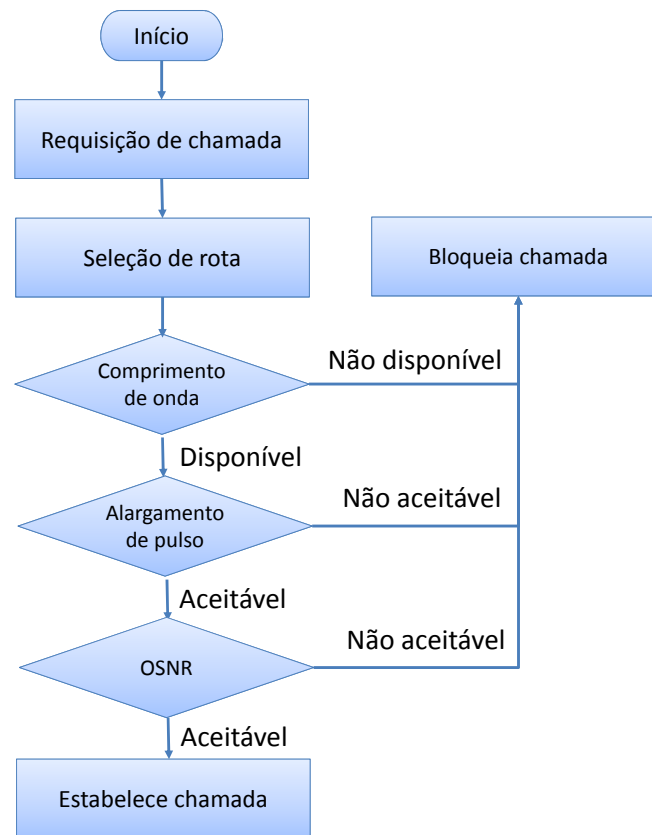


Figura 4.3: Fluxograma de execução do CAC transparente.

$$PB = \frac{C_{bloqueadas}}{C_{bloqueadas} + C_{aceitas}}, \quad (4.2)$$

em que $C_{bloqueadas}$ é a quantidade de requisições de chamadas rejeitadas pelo CAC e $C_{aceitas}$ é o número de requisições de chamadas aceitas.

4.2 Proposta de uma Nova Metodologia para Análise de Redes Ópticas

As simulações de rede realizadas por ferramentas como o SIMTON para estimativa de PB são computacionalmente caras. Em casos extremos, como na estimativa de uma rede com PB em torno de 10^{-5} , é necessário simular até alguns milhões de requisições de chamadas, e isso pode levar até minutos para ser computado. Nesta seção é proposto o uso de Redes Neurais Artificiais (RNA) para prever a PB de uma rede transparente sujeita a tráfego dinâmico. A proposta usa métricas topológicas e características da camada física.

4.2.1 Descrição

Nesta Tese é proposta uma metodologia para obtenção de um estimador rápido e preciso para PB de redes ópticas baseado em um modelo de RNA para regressão que considera as propriedades topológicas da rede, parâmetros da camada física, o algoritmo de RWA e informações sobre o tráfego da rede. O fluxograma apresentado na Figura 4.4 resume a metodologia proposta para obtenção do estimador.

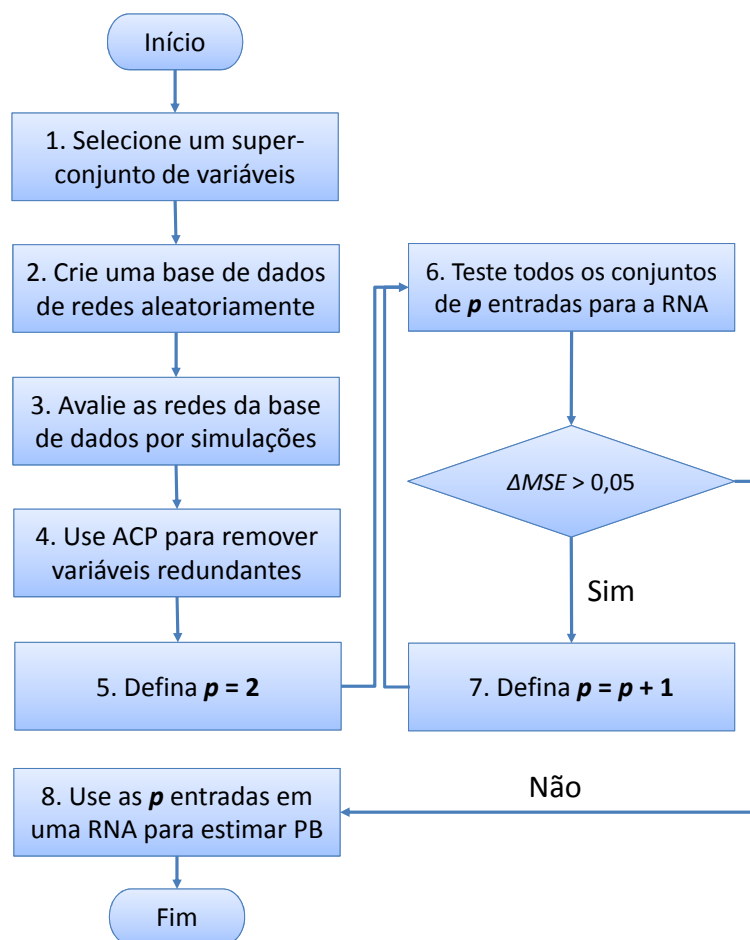


Figura 4.4: Fluxograma com os passos da metodologia proposta nesta Tese para obtenção de um estimador de PB baseado em RNA.

De acordo com a Figura 4.4, o primeiro passo consiste na seleção de um superconjunto de variáveis independentes a serem usadas como variáveis candidatas para estimativa de PB. Em seguida, é criada uma base de dados de redes ópticas por meio de um processo aleatório de geração de redes e as redes criadas são avaliadas por meio de um simulador de redes ópticas capaz de fornecer uma estimativa de PB de referência. A Análise de Componentes Principais (ACP) é usada como primeiro procedimento para reduzir o superconjunto de variáveis inde-

pendentes. ACP identifica um conjunto de variáveis independentes que podem ser usadas para representar variação nos dados analisados por meio de uma combinação linear. Assim, ACP sugere variáveis redundantes que podem ser removidas do super-conjunto inicial sem perda significativa de informação no processo de estimativa de PB. O conjunto final de variáveis de entrada pode ser obtido usando o procedimento *best-selection*, que consiste na busca pelo melhor subconjunto de p variáveis, após a análise de todas as possibilidades. O procedimento é iniciado com o teste do melhor conjunto de duas variáveis de entrada ($p = 2$). Em seguida, são testadas todas as combinações de três variáveis de entrada ($p = 3$). É calculada a diferença entre o erro médio quadrático obtido para o melhor subconjunto de p variáveis e o melhor subconjunto de $p - 1$ variáveis, designada por ΔMSE . O processo é repetido até que a diferença entre os erros fornecidos por dois melhores subconjuntos consecutivos seja inferior a 5% ($\Delta MSE < 0,05$). O processo é interrompido neste ponto e um estimador com p variáveis de entrada é obtido.

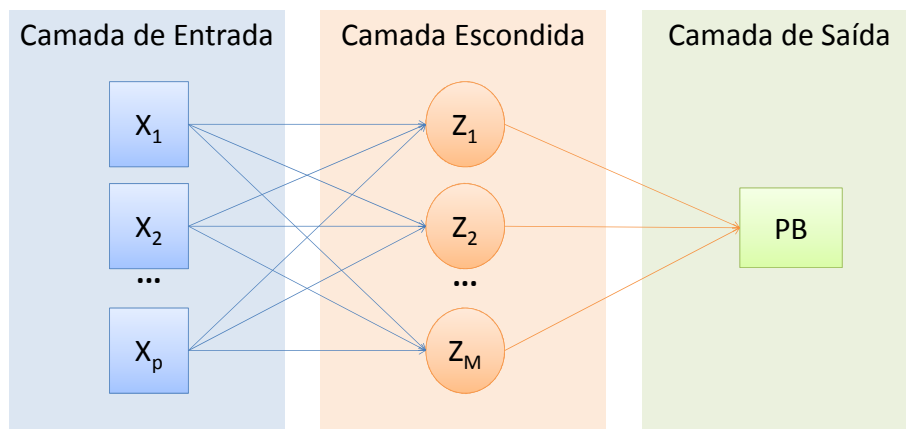


Figura 4.5: RNA com uma camada escondida para obtenção de PB.

Nesta Tese foi adotado um modelo de RNA amplamente conhecido na literatura, denominado Perceptron Multi-Camadas (*Multi-Layer Perceptron* - MLP). Uma MLP genérica, usada para regressão, é ilustrada na Figura 4.5. Neste esquema, características derivadas Z_m são criadas pela combinação das entradas X_p , e então a PB é modelada como uma combinação destas características derivadas. As características derivadas são avaliadas de acordo com

$$Z_m = \sigma\left(\alpha_0 + \sum_{p=1}^P \alpha_p X_p\right), \quad (4.3)$$

em que $m \in \{1, 2, \dots, M\}$ e M é o número de neurônios na camada escondida. α_0 é o *bias* da RNA. Cada neurônio faz uma combinação linear das entradas de acordo com uma função de ativação $\sigma(v)$. Neste trabalho é usada a função logística (sigmóide) como função de ativação, ou seja, $\sigma(v) = 1/(1 + e^{-v})$.

Portanto, PB pode ser obtida a partir das características derivadas por

$$PB = \sigma(\beta_0 + \sum_{m=1}^M \beta_m Z_m), \quad (4.4)$$

em que $Z = (Z_1, Z_2, \dots, Z_M)$ são as saídas dos neurônios da camada escondida.

Uma visão geral da arquitetura completa da proposta é ilustrada na Figura 4.6. A proposta consiste em três grandes módulos. O primeiro módulo, denominado “Motor de aprendizado”, é projetado para obter os pesos da RNA com base em um banco de dados de redes ópticas previamente avaliada por um simulador de eventos discretos. O segundo módulo, denominado “Avaliador”, é projetado para avaliar uma rede óptica com base em uma RNA previamente treinada. Um módulo auxiliar, denominado “Motor de Redes Complexas”, permite o cálculo de diversas características topológicas de uma rede e é usado por ambos os módulos “Motor de aprendizado” e “Avaliador”. Na Figura 4.6, o termo “WRON” corresponde à configuração completa de uma rede óptica, incluindo a especificação da topologia física, o número de comprimentos de onda por enlace e especificação do comutador óptico utilizado.

O módulo “Motor de aprendizado” é chamado apenas uma vez, de modo a criar a base de dados de redes aleatórias e treinar a RNA usando esta base de dados. Na Figura 4.6, a caixa “Crie uma base de dados de WRONs” implica na criação de um grande número de redes ópticas usando um procedimento de geração aleatório. O conjunto de dados deve ser preferencialmente composto por redes com diferentes números de nós e posições geográfica de modo a obter um estimador que possa ser usado para diferentes cenários de avaliação de redes. O simulador de eventos discretos usado neste passo é um processo auxiliar predefinido. O módulo “Motor de aprendizado” avalia o impacto de cada variação na topologia física, no número de comprimentos de onda e na especificação de dispositivos por meio de uma estimativa de PB obtida a partir do simulador de rede. Em seguida, o módulo “Motor de aprendizado” usa esta informação para comandar o processo de aprendizagem supervisionada da RNA (passo relacionado com a caixa “Calcule os pesos da RNA”).

É importante ressaltar que a avaliação de uma rede óptica usando a RNA treinada após o processo de aprendizado é um procedimento rápido, visto que é necessário apenas calcular métricas topológicas da rede e as Equações (4.3) e (4.4). Este processo é várias ordens de grandeza mais rápido quando comparado ao uso de um simulador de eventos discretos. Além disso, se este método for considerado para uso como parte de uma ferramenta de análise de redes, o módulo “Motor de aprendizado” pode ser executado antes da fase de operação da ferramenta e o módulo “Avaliador” pode ser usado como um sistema em tempo real. Portanto, considerando a arquitetura mostrada na Figura 4.6, é possível observar que a proposta apresentada neste trabalho foi concebida para apresentar uma resposta quase instantânea.

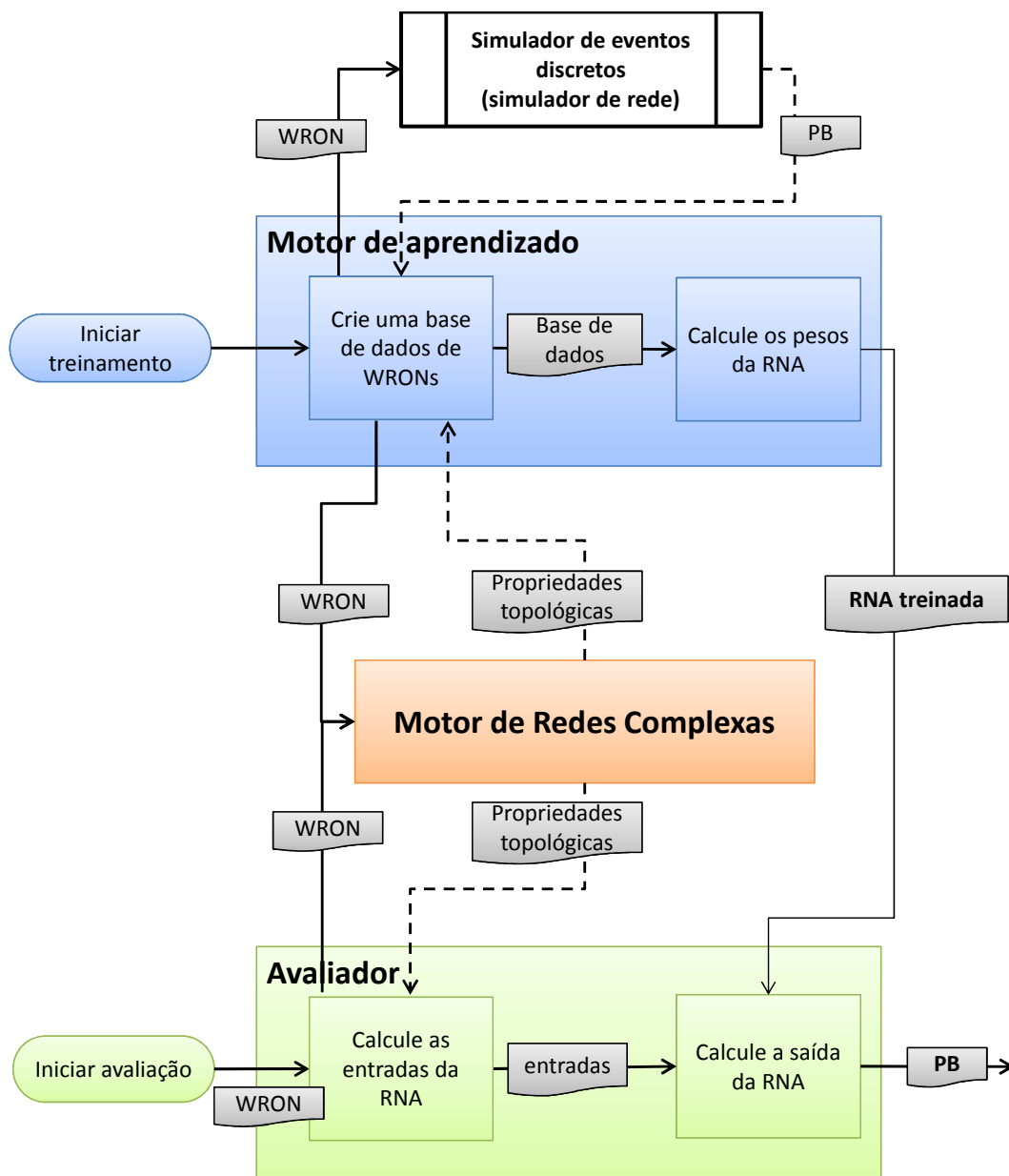


Figura 4.6: Uma visão geral do analisador de redes proposto.

4.2.2 Arranjo Experimental

A investigação sobre a viabilidade do uso de redes neurais para estimar PB de redes ópticas foi conduzida em duas etapas. O objetivo da primeira etapa foi validar a hipótese de que métricas topológicas derivadas da Ciência das Redes podem ser usadas como entrada para um motor de aprendizado para fornecer uma estimativa de PB com baixo erro. A segunda etapa foi conduzida com a finalidade de aprofundar a investigação sobre a metodologia de seleção de variáveis, sobre o desempenho da RNA, e sobre a capacidade de aplicar a proposta

para diferentes coberturas geográficas.

A primeira etapa do estudo fez uso de uma grande base de dados de configurações de redes derivadas da NSFNet (vide Figura 3.13). A NSFNet é uma rede amplamente usada na literatura e sua aplicação como estudo de caso inicial facilitou a validação dos resultados. A base de dados usada na primeira etapa contém 32.000 diferentes redes derivadas da distribuição de nós da NSFNet. Na segunda etapa foi realizada uma investigação sobre o tamanho adequado da base de treinamento e foram usadas outras distribuições geográficas para validar a proposta. Nesta segunda etapa foram usados outros 3 casos diferentes, obtidos a partir de redes de transporte publicamente disponíveis, implantadas em diferentes países. Além disso, na segunda etapa de investigação foi realizada uma análise comparativa sobre o impacto das penalidades da camada física sobre o estimador, considerando a precisão das estimativas e o tempo de processamento. Nas próximas subseções é feito um detalhamento sobre os métodos empregados para condução dos experimentos em cada uma das etapas do estudo.

Arranjo Experimental Usado na Etapa Inicial do Estudo

Nesta fase inicial foi estudado o impacto de 15 diferentes variáveis independentes para estimativa da PB de uma rede WDM. Os parâmetros estudados estão listados na Tabela 4.1. Os dois primeiros parâmetros estão associados com a camada física da rede óptica e os demais estão relacionados com características topológicas. O diâmetro da rede (X_{12}) e o comprimento médio dos caminhos (X_{13}), ambos medidos em quilômetros, são indicadores híbridos, pois combinam informações da topologia e a distância física (que por sua vez interfere no cálculo das penalidades). Esta Tese usa como premissa o fato de que um algoritmo de RWA será configurado no simulador de redes e a RNA irá aprender o impacto que este RWA em específico fornece sobre as estimativas de PB. Portanto, o algoritmo de RWA não é uma informação explícita para a RNA, mas implicitamente a sua escolha afeta no treinamento e na resposta da RNA. Se uma arquitetura genérica exigir uma informação explícita sobre o algoritmo de RWA, é possível adaptar a proposta para que a RNA possua uma nova entrada no formato de um rótulo para um algoritmo de RWA em específico. Nesta abordagem, amostras de redes ópticas com PB avaliadas para os diferentes tipos de RWA devem ser fornecidas, de modo a viabilizar o processo de aprendizado.

Foi realizado um procedimento formal de subseleção de variáveis de modo a encontrar o melhor subconjunto com p parâmetros considerando o compromisso entre precisão, generalização e facilidade de interpretação.

Para simulação das redes ópticas foi usado o simulador SIMTON [11]. Todas as redes foram avaliadas pelo SIMTON para uma carga de 200 erlangs de modo a obter a PB usada como referência para a RNA. Em seguida, esta base de dados foi avaliada novamente para uma carga de 100 erlangs com o objetivo de permitir a generalização desta proposta com relação a mudanças na carga de tráfego. O valor máximo de carga foi definido com o objetivo de permitir

Tabela 4.1: Variáveis independentes consideradas para estimativa de PB.

Índice	Variável	Definição
X_1	W	Número de comprimentos de onda
X_2	δ	Fator de isolamento das portas do ROADM
X_3	CC	Coefficiente de agrupamento
X_4	q	Densidade
X_5	$I(\hat{\mathcal{F}})$	Entropia da DFT dos autovalores do Laplaciano
X_6	AC	Conectividade algébrica
X_7	NC	Conectividade natural
X_8	\hat{k}	Grau médio
X_9	\bar{c}	Comprimento médio dos caminhos (saltos)
X_{10}	D	Diâmetro
X_{11}	$I(G)$	Entropia
X_{12}	D_{km}	Diâmetro (em km)
X_{13}	\bar{c}_{km}	Comprimento médio dos caminhos (em km)
X_{14}	ρ	Raio espectral
X_{15}	CR	Concentração de rotas

obter topologias de redes que apresentem $PB < 10^{-3}$. O SIMTON foi configurado para parar as simulações após a ocorrência de 1.000 bloqueios de chamadas ou quando um máximo de 10^7 chamadas for alcançado. A Qualidade de Transmissão (QoT) de cada caminho óptico foi avaliado considerando os seguintes efeitos da camada física: emissão espontânea amplificada (*Amplified Spontaneous Emission* - ASE) dos amplificadores, o efeito de saturação de ganho dos amplificadores, a dispersão cromática residual, o *crosstalk* nos ROADMs e o efeito da mistura de quatro ondas (*Four Wave Mixing* - FWM). Os parâmetros usados nas simulações de redes WDM estão listados na Tabela 4.2. Vale a pena ressaltar que outro simulador de redes diferente do SIMTON, que considere penalidades da camada física, poderia ser usado para funcionar como “supervisor” da RNA.

Foi definido um conjunto com 12 classes de intervalos distintos para PB na base de dados com o objetivo de melhorar o desempenho do treinamento da RNA. Uma tentativa com 4 classes apresentou uma qualidade inferior quando comparada com o uso de 12 classes [92]. Foi executado um procedimento de pré-seleção na base de dados original, de modo a garantir que haja pelo menos 1.500 redes em cada uma das 12 classes. A Tabela 4.3 apresenta as classes que foram definidas. Portanto, foi utilizada uma base de dados contendo 18.000 redes WDM para cada carga de tráfego. Esta base de dados foi dividida em 3 diferentes grupos para serem usados para o treinamento, validação e teste da RNA. Foi utilizado um percentual de 50% das redes para treinamento, 25% para validação e 25% para testes.

De modo a obter o melhor subconjunto de variáveis, inicialmente foi usada a Análise de Componentes Principais (ACP) para encontrar variáveis independentes redundantes. Em seguida, tentou-se aplicar Regressão Linear Múltipla (RLM) [93] como uma primeira aproximação para o estimador de PB. Foi usado o erro médio quadrático (MSE - *Mean Squared Error*) para aferição da qualidade do estimador, de acordo com a Eq. (4.5).

Tabela 4.2: Parâmetros de simulação das redes WDM.

Parâmetro	Valor	Definição
α	0,2 dB/km	Coefficiente de atenuação da fibra
Δt_{QoT}	0,10	Alargamento máximo permitido para o pulso óptico
Δf	100 GHz	Espaçamento entre canais
λ_i	1528,77 nm	Primeiro comprimento de onda da grade
λ_0	1450 nm	Zero de dispersão na fibra de transmissão
λ_{0RD}	1528,77 nm	Zero de dispersão residual
B_o	100 GHz	Largura de banda do filtro
B	10 Gb/s	Taxa de transmissão
D_{DCF}	-110 ps/km.nm	Coefficiente de dispersão da fibra de compensação
D_{PMD}	0,04 ps/ \sqrt{km}	Coefficiente de PMD
D_{TX}	4,5 ps/km.nm	Coefficiente de dispersão da fibra de transmissão
L_{Mux}	3 dB	Perda no multiplexador
L_{Demux}	3 dB	Perda no demultiplexador
L_{Switch}	3 dB	Perda no comutador óptico
NF	5 dB	Figura de ruído do amplificador
P_{Sat}	20 dB	Potência de saturação do amplificador
P_{TX}	3 dBm	Potência de transmissão
S_{DCF}	-1,87 ps/km.nm ²	<i>Slope</i> da fibra de compensação
S_{TX}	0,045 ps/km.nm ²	<i>Slope</i> da fibra de transmissão
$OSNR_{in}$	30 dB	Relação sinal-ruído óptica na entrada
$OSNR_{QoS}$	17 dB	OSNR mínima no receptor
W_{MIN}	4	Número mínimo de comprimentos de onda por enlace
W_{MAX}	40	Número máximo de comprimentos de onda por enlace

Tabela 4.3: Classes de intervalos de PB para criar a base de dados de redes ópticas.

Classe	Limite inferior	Limite superior
1	1E-5	1E-3
2	1E-3	0,01
3	0,01	0,10
4	0,10	0,20
5	0,20	0,30
6	0,30	0,40
7	0,40	0,50
8	0,50	0,60
9	0,60	0,70
10	0,70	0,80
11	0,80	0,90
12	0,90	1,00

$$MSE = RSS(\beta)/N, \quad (4.5)$$

em que N é o número de redes usadas para o cálculo do erro médio e $RSS(\beta)$ (*Residual Sum of Squares*) é o erro residual do vetor de pesos β , baseado na soma dos quadrados. Ou seja, $RSS(\beta) = \sum_{i=1}^N (y_i - x_i^T \beta)^2$, em que y_i é o valor esperado de PB para a rede i e $x_i^T \beta$ é o valor de PB fornecido pelo estimador linear.

O processo é iniciado com as duas variáveis mais importantes (de acordo com a análise de correlação) e em seguida é avaliado o melhor modelo com $k \in \{2, \dots, p\}$ parâmetros para

seleção de entradas para a RLM. Estudos demonstram que a inclusão de novas variáveis deve respeitar um compromisso entre o aumento da precisão do estimador e a capacidade de generalizar novos casos [93]. Se o processo de seleção for configurado para que o erro médio diminua para valores muito próximos de zero, a capacidade de generalização do modelo é comprometida. Portanto, nesta Tese adotou-se $\Delta MSE < 0,05$ como critério de parada para o processo de seleção de variáveis. De acordo com os resultados obtidos, RLM não é adequado para o problema considerado, pois foi obtido $MSE > 10^{-3}$ para todas as tentativas no caso linear. Contudo, o uso de RLM foi útil para o entendimento da relação entre PB e as variáveis independentes quando atuam em conjunto.

Em seguida, foi adotada uma RNA para estimativa de PB. A camada de entrada foi definida usando o mesmo procedimento descrito para RLM. Foi usado o algoritmo de Levenberg-Marquardt (LM) para treinamento da RNA. O método LM é uma modificação do método de Gauss-Newton e consiste em uma minimização iterativa pelo método dos mínimos quadrados. Em muitas classes de problemas, o método LM possui desempenho superior aos métodos tradicionais de gradiente descendente, pois a regra de atualização leva em consideração tanto a inclinação da superfície do erro (método do gradiente descendente) quanto a curvatura desta superfície (método de Gauss-Newton) [94]. A implementação do algoritmo finaliza as iterações de LM quando o erro de validação começa a aumentar (validação cruzada) ou quando o número de iterações alcança $I_{max} = 500,000$. I_{max} foi escolhido de modo a ser um valor ligeiramente superior ao valor médio de ocorrência da validação cruzada, com o objetivo de evitar que o processo execute indefinidamente quando estiver estagnado em mínimos locais. Todos os pesos da RNA foram ajustados para variarem no intervalo $[0, 1]$. Foram usados $M = 2p + 1$ neurônios na camada escondida da RNA.

Todos os experimentos foram realizados em um computador com processador Intel i5 com 4 GB de RAM. Todos os módulos da Figura 4.6 foram implementados usando a linguagem de programação Java. Operações básicas de álgebra linear foram suportadas pela biblioteca Apache Commons Math API [74].

Arranjo Experimental Usado na Etapa Final do Estudo

Esta etapa do estudo aprofunda a avaliação do compromisso entre a precisão na estimativa de PB e o tempo de processamento requerido pela abordagem proposta. Serão considerados nesta etapa duas diferentes modelagens para o problema de regressão:

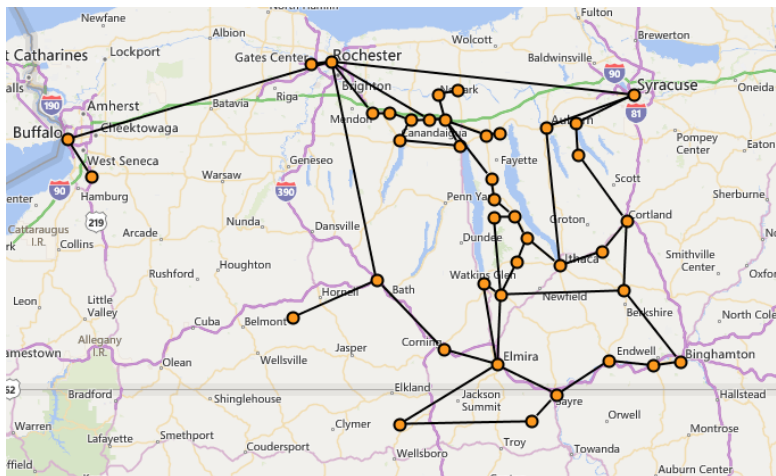
- **Problema Tipo A:** considera a topologia de fibra de uma rede óptica, o número máximo de comprimentos de onda disponíveis e as penalidades da camada física. A QoT de cada caminho óptico é avaliado considerando ASE, dispersão cromática residual e o crosstalk nos ROADMs.
- **Problema Tipo B:** considera apenas a topologia de fibra e o número máximo de

comprimentos de onda. A estimativa de PB leva em consideração apenas bloqueio devido à falta de comprimentos de onda para atender às requisições de chamadas.

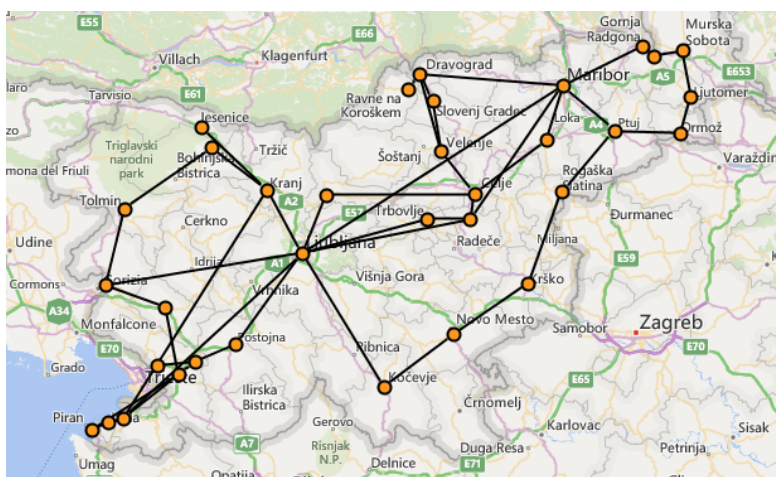
Com o objetivo de avaliar se a abordagem baseada em RNA oferece desempenho semelhante quando aplicada a diferentes distribuições geográficas, foram escolhidos mais três cenários para validação da proposta. A Figura 4.7 ilustra a topologia de fibra dos novos casos considerados para estudo. Os dois principais critérios adotados para escolha das redes foram: oferecer publicamente os detalhes sobre a topologia física das redes e possuir dimensões geográficas compatíveis com uma rede óptica transparente. Para especificação do segundo critério foi considerada a restrição de diâmetro físico limitado em 400 quilômetros, ou seja, foram procuradas redes em operação nas quais os dois nós mais afastados possuíssem uma distância máxima de 400 quilômetros. Apesar de existirem diversas redes de fibra óptica comerciais em operação atualmente, a disponibilidade de informações sobre estas redes ainda é escassa. Além disso, evitou-se no estudo a seleção de redes muito semelhantes, como redes com uma mesma quantidade de nós ou localizadas em regiões próximas.

A rede FLTG [95] é uma rede de transporte americana que possui 41 nós e fica localizada na região conhecida como “Finger Lakes”. FLTG cobre geograficamente Rochester, Ithaca e Syracuse. A rede FLTG possui mais de 3200 quilômetros de fibra óptica e é a maior rede considerada neste estudo em termos de número de nós e de abrangência geográfica. ARNES [96] possui 34 nós e fornece uma infraestrutura de rede de transporte óptica para Eslovênia. As conexões oferecidas pela rede ARNES são em sua maioria de 1 Gbit/s ou de 10 Gbit/s, mas também estão disponíveis algumas poucas conexões de 40 Gbit/s para interligação com a rede europeia GÉANT. Dentre as três redes escolhidas, ARNES é uma rede intermediária em termos de número de nós e cobertura geográfica. MEDIANET [97] é uma rede de transporte de 18 nós do ISP (*Internet Service Provider*) HSE. HSE fornece serviços de Internet e de voz no sul de Hesse, Alemanha. MEDIANET possui 2.000 quilômetros de enlaces de fibra óptica que oferecem largura de banda de até 10 Gbit/s. MEDIANET é a menor rede considerada no estudo em termos de número de nós e de cobertura geográfica. Vale a pena enfatizar que as análises que serão realizadas individualmente sobre cada um dos três casos consiste no uso da localização geográfica dos nós destas redes para geração e avaliação de novas topologias físicas. Portanto, as topologias de fibra apresentadas na Fig. 4.7 servem apenas para ilustrar a cobertura geográfica dada pela posição dos nós e a distância relativa entre estes nós, uma vez que os enlaces apresentados nesta figura representam apenas uma dentre muitas possibilidades de interconexão entre os nós.

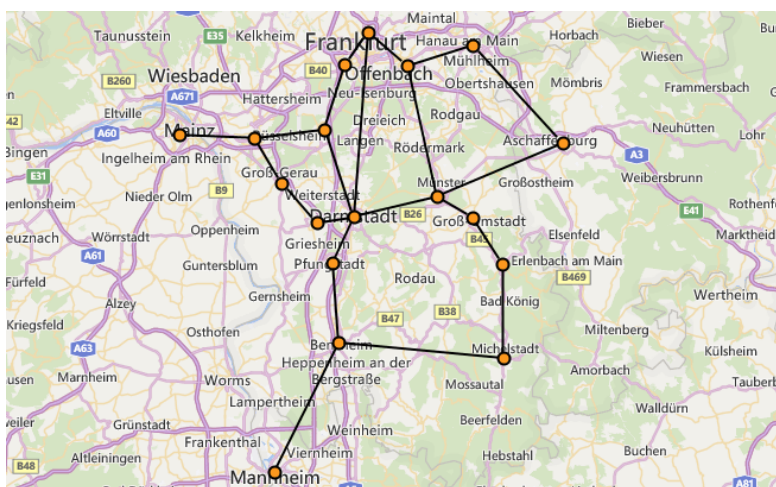
Nesta fase final foi estudado o impacto de 20 diferentes variáveis independentes para estimativa da PB de uma rede WDM. A Tabela 4.4 apresenta todas as variáveis analisadas, incluindo uma coluna referente à complexidade computacional de cada métrica. Destas variáveis, 16 foram descritas na Subseção 4.2.2. As últimas quatro variáveis foram incluídas na tentativa de melhorar a avaliação de penalidades da camada física. As três primeiras variáveis



(a) Rede FLTG (EUA), com 41 nós.



(b) Rede ARNES (Eslovênia), com 34 nós.



(c) Rede MEDIANET (Alemanha), com 18 nós.

Figura 4.7: Topologias de fibra de três diferentes redes de transporte.

da Tabela 4.4 correspondem a valores predefinidos e, portanto, não precisam ser calculadas. As demais variáveis são calculadas, mas a complexidade de todas as métricas é de tempo polinomial, variando entre $\mathcal{O}(n^2)$ e $\mathcal{O}(n^4)$. Contudo, considerando que o valor de n para redes de transporte é tipicamente pequeno, a variação do custo computacional das métricas não pode ser negligenciada.

A margem de OSNR (ΔOSNR) é uma média da diferença entre o limiar de OSNR de recepção e a OSNR recebida, calculada sobre todos os caminhos ópticos. O desvio padrão de ΔOSNR mede a variação da margem de OSNR considerando todos os caminhos ópticos. Estas duas métricas, assim como CR, possui complexidade $\mathcal{O}(n^4)$ pois o fluxo dominante da métrica está relacionado com o uso de um algoritmo de menor caminho como o algoritmo de Dijkstra [98], pois cada rota encontrada é usada nos cálculos adicionais da métrica. O desvio padrão dos caminhos mínimos mede a variação dos comprimentos dos caminhos mínimos de todos os caminhos ópticos, em quilômetros. Esta métrica, assim como \bar{c} e \bar{c}_{km} , possui complexidade $\mathcal{O}(n^3)$ pois o fluxo dominante da métrica está relacionado com o uso de um algoritmo de menor caminho como o algoritmo de Floyd-Warshall [98], no qual o comprimento total de cada caminho é obtido, mas não há interesse na sequência de nós usados em cada rota. A densidade de fibras implantadas corresponde à taxa entre o número de quilômetros de fibras implantadas na rede e o número máximo de fibras implantadas, correspondente à uma rede completamente conectada. Esta métrica fornece uma informação adicional sobre o conceito tradicional de densidade de enlaces e pode melhorar a estimativa quando estão presentes penalidades da camada física relacionadas com distância geográfica. A complexidade desta última métrica é $\mathcal{O}(n^2)$, pois está relacionada com todas as possibilidades de enlaces de fibra óptica em uma rede contendo n nós.

O tempo necessário para a finalização do método *best-subset selection* é impraticável para 20 diferentes variáveis e para três diferentes cenários. Há um total de $C_{20}^p = \frac{20!}{p!(20-p)!}$ combinações para cada valor de p variáveis. Considerando as três redes e modelos de RNA com no máximo 6 variáveis ($p \in \{2, 3, 4, 5, 6\}$), há um total de 181317 combinações. Além disso, cada combinação é testada mais de uma vez para avaliação da variabilidade dos dados no treinamento da RNA. Considerando que o método *best-subset selection* usado para a análise dos resultados da etapa inicial comportou-se de forma semelhante ao método *forward-stepwise*, nesta etapa foi feito uso deste último método devido à sua simplicidade.

Nesta etapa, também foi usado o simulador SIMTON [11] para fornecimento das estimativas de referência de PB e para o problema Tipo A, adotou-se a mesma configuração usada na subseção 4.2.2, excetuando-se a avaliação do efeito FWM. Para o problema Tipo B, as penalidades da camada física foram desconsideradas. Além disso, com base no estudo inicial, percebeu-se que são poucas as configurações de redes conectadas que possuem PB acima de 0,70 e nesta fase final adotou-se um agrupamento de configurações de redes com apenas 10 classes, conforme a Tabela 4.5. As redes MEDIANET foram submetidas a uma carga de 100,

Tabela 4.4: Variáveis independentes incluídas na etapa final do estudo sobre a estimativa de PB.

Índice	Variável	Definição	Complexidade
X_0	L	Carga da rede, em erlangs	-
X_1	W	Número de comprimentos de onda	-
X_2	δ	Fator de isolamento das portas do ROADM	-
X_3	CC	Coefficiente de agrupamento	$\mathcal{O}(n^3)$
X_4	q	Densidade	$\mathcal{O}(n^2)$
X_5	$I(\hat{\mathcal{F}})$	Entropia da DFT dos autovalores do Laplaciano	$\mathcal{O}(n^3)$
X_6	AC	Conectividade algébrica	$\mathcal{O}(n^3)$
X_7	NC	Conectividade natural	$\mathcal{O}(n^3)$
X_8	\hat{k}	Grau médio	$\mathcal{O}(n^2)$
X_9	\bar{c}	Comprimento médio dos caminhos (saltos)	$\mathcal{O}(n^3)$
X_{10}	D	Diâmetro	$\mathcal{O}(n^3)$
X_{11}	$I(G)$	Entropia	$\mathcal{O}(n^2)$
X_{12}	D_{km}	Diâmetro (em km)	$\mathcal{O}(n^3)$
X_{13}	\bar{c}_{km}	Comprimento médio dos caminhos (em km)	$\mathcal{O}(n^3)$
X_{14}	ρ	Raio espectral	$\mathcal{O}(n^3)$
X_{15}	CR	Concentração de rotas	$\mathcal{O}(n^4)$
X_{16}	$\overline{\Delta OSNR}$	Margem média de OSNR	$\mathcal{O}(n^4)$
X_{17}	$\sigma_{\Delta OSNR}$	Desvio padrão de $\Delta OSNR$	$\mathcal{O}(n^4)$
X_{18}	σ_{PL}	Desvio padrão dos caminhos mínimos	$\mathcal{O}(n^3)$
X_{19}	q_{km}	Densidade de implantação de fibras	$\mathcal{O}(n^2)$

150 e 200 erlangs. As redes ARNES e FLTG foram submetidas a uma carga de 100, 200 e 300 erlangs.

Finalmente, é importante realizar uma ressalva sobre a metodologia empregada neste estudo e sobre a validade das conclusões obtidas. Para cada uma das três distribuições geográficas estudadas foi feito um estudo abrangente considerando milhares de diferentes configurações de rede. Portanto, conclusões obtidas sobre os casos individuais, tais como o melhor conjunto de variáveis para representar a PB ou o número mínimo de redes requerido para treinamento, de fato são estatisticamente confiáveis no contexto de cada cenário avaliado. Contudo, análises sobre o impacto de diferentes coberturas geográficas sobre o desempenho da RNA serão feitas de forma qualitativa, pois o uso de apenas três casos não permite uma generalização estatisticamente confiável para todos os possíveis cenários.

4.2.3 Resultados Obtidos

Discussão dos Resultados da Etapa Inicial do Estudo

A Tabela 4.6 apresenta a matriz de correlação (R) das variáveis independentes, calculada usando o coeficiente de correlação de Pearson. R fornece um primeiro passo na direção da escolha de um conjunto apropriado de variáveis. Por exemplo, é possível observar que X_4 (densidade) e X_8 (grau médio) apresentam um índice de correlação igual a 1,00. Portanto, deve-se usar apenas a primeira ou a segunda variável. Além disso, os parâmetros relacionados com a camada óptica, como X_1 (número de comprimentos de onda) e X_2 (fator de isolamento do

Tabela 4.5: Classes de intervalos de PB para criar a base de dados de redes ópticas na etapa final do estudo.

Classe	Limite inferior	Limite superior
1	1E-5	1E-3
2	1E-3	0,01
3	0,01	0,10
4	0,10	0,20
5	0,20	0,30
6	0,30	0,40
7	0,40	0,50
8	0,50	0,60
9	0,60	0,70
10	0,70	1,00

ROADM) não estão diretamente correlacionados com os parâmetros topológicos. Portanto, estes parâmetros devem ser incluídos no modelo de modo a melhorar a precisão do estimador de PB. O próximo passo é calcular os autovalores e os autovetores de R de modo a proceder com a análise de componentes principais.

Tabela 4.6: Matriz de correlação das variáveis independentes (vide Tabela 4.1).

	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}
X_1	0,41	0,29	0,32	-0,32	0,30	0,30	0,32	-0,29	-0,24	-0,08	-0,44	-0,46	-0,15	-0,43
X_2	-	0,19	0,20	-0,18	0,20	0,21	0,20	-0,12	-0,07	-0,08	-0,30	-0,34	-0,02	-0,18
X_3	-	-	0,93	-0,86	0,87	0,93	0,93	-0,81	-0,72	0,05	-0,40	-0,61	-0,21	-0,59
X_4	-	-	-	-0,95	0,97	1,00	1,00	-0,85	-0,75	-0,07	-0,42	-0,64	-0,29	-0,62
X_5	-	-	-	-	-0,92	-0,93	-0,95	0,86	0,78	0,11	0,43	0,62	0,36	0,64
X_6	-	-	-	-	-	0,97	0,97	-0,74	-0,64	-0,24	-0,37	-0,58	-0,15	-0,50
X_7	-	-	-	-	-	-	1,00	-0,80	-0,70	-0,08	-0,40	-0,62	-0,23	-0,57
X_8	-	-	-	-	-	-	-	-0,85	-0,75	-0,07	-0,42	-0,64	-0,29	-0,62
X_9	-	-	-	-	-	-	-	-	0,95	-0,22	0,42	0,61	0,59	0,77
X_{10}	-	-	-	-	-	-	-	-	-	-0,23	0,36	0,52	0,60	0,71
X_{11}	-	-	-	-	-	-	-	-	-	-	0,01	-0,00	-0,37	-0,20
X_{12}	-	-	-	-	-	-	-	-	-	-	-	0,83	0,24	0,54
X_{13}	-	-	-	-	-	-	-	-	-	-	-	-	0,29	0,64
X_{14}	-	-	-	-	-	-	-	-	-	-	-	-	-	0,55

A Tabela 4.7 apresenta os autovalores da matriz R (Tabela 4.6). De acordo com a Tabela 4.7, as primeiras três principais componentes explicam mais de 80% da variação total nos dados. Portanto, serão estudados apenas os autovetores relacionados com estes três principais componentes.

A Tabela 4.8 apresenta os autovetores normalizados da matriz de correlação (cada autovetor foi dividido por $\sqrt{\lambda_i}$). A primeira coluna da Tabela 4.8 é o autovetor relacionado com o maior autovalor, a segunda coluna é o autovetor relacionado com o segundo maior autovalor e a terceira coluna é o autovetor relacionado com o terceiro maior autovalor. O primeiro componente principal (PC-1) ignora completamente apenas X_{11} (entropia) de modo a explicar a variação nos dados. Se forem consideradas apenas variáveis com correlação acima de 0,9 em PC-1, obtém-se as variáveis X_3 , X_4 , X_5 , X_6 , X_7 , X_8 e X_9 . É possível observar que as variáveis relacionadas com a topologia de fibra são muito importantes para explicar a variação nos dados das variáveis independentes. Também é possível concluir, que de acordo com PC-1, X_5 e X_9 aumentam de valor se X_3 , X_4 , X_6 , X_7 e X_8 diminuem. Por exemplo, um aumento

Tabela 4.7: Percentual da variação total nos dados explicada pelas componentes principais da matriz de correlação de variáveis que influenciam a PB.

Componente Principal	Autovalor	% Variação	% Variação Acumulada
1	8,72	58,13	58,13
2	1,82	12,15	70,27
3	1,66	11,04	81,31
4	0,77	5,12	86,43
5	0,66	4,42	90,84
6	0,54	3,58	94,43
7	0,29	1,92	96,35
8	0,22	1,47	97,82
9	0,13	0,86	98,68
10	0,10	0,65	99,33
11	0,06	0,37	99,70
12	0,03	0,19	99,89
13	0,01	0,10	99,99
14	0,00	0,01	100,00
15	0,00	0,00	100,00

em densidade irá criar rotas alternativas e \bar{c} irá diminuir; e de acordo com o Capítulo 3, $I(\hat{\mathcal{F}})$ versus d é uma curva monotonicamente decrescente. A mesma interpretação pode ser dada a outras variáveis de PC-1. PC-2 tenta explicar o contraste entre valores de X_{11} (entropia) e X_{14} (raio espectral). PC-3 fornece informação sobre uma diminuição em X_{12} (diâmetro) e um aumento em X_1 (número de comprimentos de onda) e X_2 (qualidade do ROADM). A análise de componentes principais é útil como um primeiro passo em direção a um motor de regressão. Contudo, o melhor modelo linear encontrado apresenta coeficiente de determinação $R^2 = 0,89$ e $\overline{MSE} = 9,28 \cdot 10^{-3}$. Um estimador para PB que apresenta $\overline{MSE} = 9,28 \cdot 10^{-3}$ não é adequado para o problema de análise de redes ópticas, pois redes WDM reais com PB $< 10^{-2}$ são muito comuns. Portanto, na próxima seção serão analisados os resultados que foram obtidos por um modelo de regressão baseado em RNA.

Em seguida, este estudo aplicou o procedimento de melhor seleção de modo a encontrar a camada de entrada para a RNA que possui o menor número de variáveis sem, no entanto, perder a precisão na previsão da resposta. O critério de parada é um aumento em precisão inferior a 0,05 quando uma nova variável $p + 1$ é incluída sobre um modelo que possui p variáveis ($\Delta \overline{MSE} < 0,05$).

Foram realizadas 31 execuções para cada conjunto de p variáveis de entrada. A Figura 4.8 apresenta *box-plots* para o melhor \overline{MSE} (valor médio do erro médio quadrático) obtido para cada valor de p . A Tabela 4.9 fornece o valor médio, o desvio padrão, a melhoria sobre o erro e o tempo de treinamento (em minutos) para cada melhor subconjunto de p entradas. É importante ressaltar que cada linha da Tabela 4.9 representa a melhor escolha com p parâmetros, pois todas as combinações de p entradas foram avaliadas. Este estudo indica que não há vantagem em escolher mais do que sete parâmetros de entrada, pois o valor de

Tabela 4.8: Correlação entre componentes principais e variáveis originais para o problema de regressão de PB.

Variável	PC-1	PC-2	PC-3
X_1	-0,43	0,08	-0,64
X_2	-0,26	0,22	-0,62
X_3	-0,91	0,13	0,17
X_4	-0,96	0,18	0,18
X_5	0,95	-0,12	-0,16
X_6	-0,90	0,36	0,19
X_7	-0,94	0,23	0,19
X_8	-0,96	0,18	0,18
X_9	0,92	0,28	-0,15
X_{10}	0,84	0,35	-0,18
X_{11}	-0,02	-0,79	0,06
X_{12}	0,57	0,05	0,61
X_{13}	0,76	0,01	0,46
X_{14}	0,43	0,72	0,02
X_{15}	0,76	0,37	0,20

\overline{MSE} diminui apenas 0,03 quando a sétima variável é incluída e o limiar definido para o procedimento foi de 0,05.

Tabela 4.9: Melhores modelos de regressão com p variáveis para estimativa de PB para a rede NSFNet.

p	Camada de entrada	\overline{MSE}	σ_{MSE}	$\Delta\overline{MSE}$	Tempo (min.)
2	W, δ	4,02E-3	4,6E-5	-	57
3	W, δ, q	5,62E-4	2,4E-4	0,86	70
4	W, δ, q, CR	3,23E-4	1,6E-5	0,43	82
5	W, δ, q, CR, CC	2,83E-4	1,5E-5	0,12	93
6	$W, \delta, q, CR, CC, I(\hat{\mathcal{F}})$	2,66E-4	6,8E-6	0,06	104
7	$W, \delta, q, CR, CC, I(\hat{\mathcal{F}}), \bar{c}(km)$	2,57E-4	9,7E-6	0,03	118

Na Figura 4.8, “SIMTON A” é uma configuração do simulador de redes que para o SIMTON após 500 bloqueios de chamadas ou quando um máximo de 10^5 chamadas é atingido e “SIMTON B” é uma configuração que para o SIMTON após 1.000 bloqueios ou quando um máximo de 10^7 chamadas é simulado. Portanto, de acordo com a Figura 4.8, a \overline{MSE} obtida para os simuladores de eventos discretos é apenas um pouco melhor do que a proposta deste trabalho quando “SIMTON B” é usado e “SIMTON A” tem pior desempenho do que esta nova proposta. Uma análise detalhada sobre a base de dados usada nos experimentos demonstrou que o modelo com 7 parâmetros fornece uma MSE abaixo de $1,7 \cdot 10^{-4}$ para redes com PB no intervalo de $10^{-4} < PB < 0,9$.

A Figura 4.9 ilustra duas redes ópticas, a e b , avaliadas por ambos SIMTON e a RNA

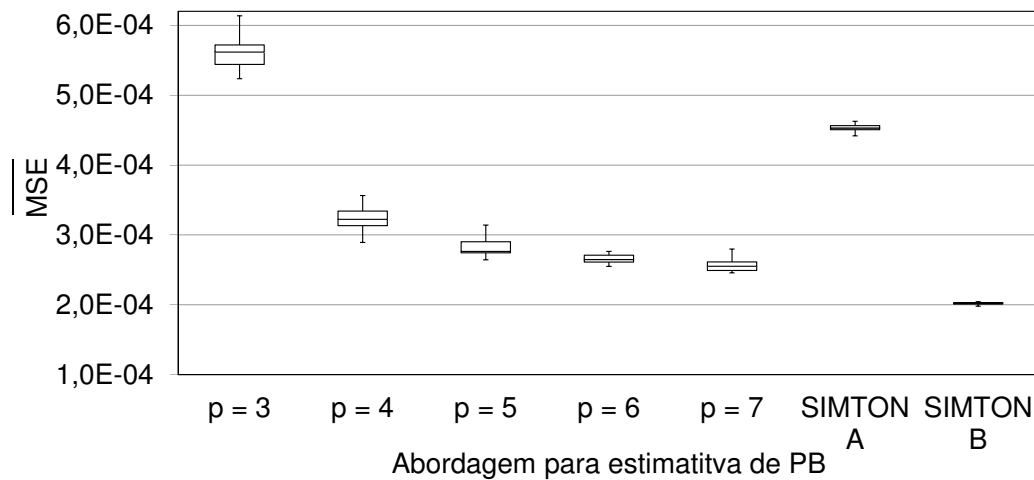


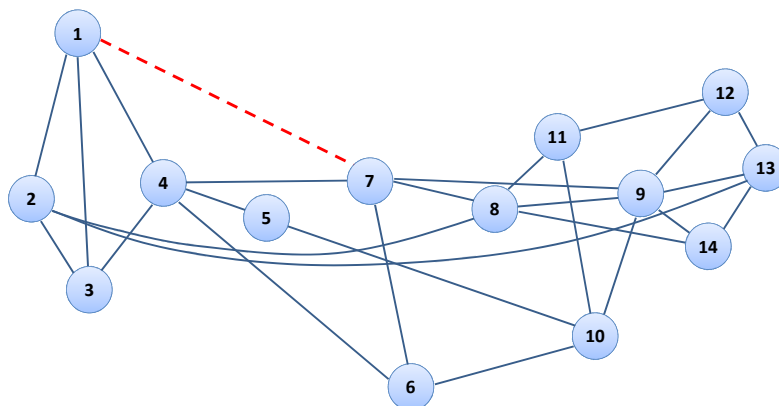
Figura 4.8: \overline{MSE} de PB para os “melhores modelos” usando diferentes números de (p) entradas e dados do SIMTON A e B, para uma rede sujeita a um tráfego de 200 erlangs.

com $p = 7$, para uma carga de 200 erlangs. Estas redes são muito semelhantes, exceto pelos enlaces destacados. SIMTON fornece $PB_a = 6,33 \cdot 10^{-3}$ e $PB_b = 2,41 \cdot 10^{-2}$, para as redes da Figura 4.9(a) e da Figura 4.9(b), respectivamente, ao passo que a RNA fornece $PB_a = 2,12 \cdot 10^{-3}$ e $PB_b = 1,42 \cdot 10^{-2}$. Para este exemplo, $MSE = 5,79 \cdot 10^{-5}$.

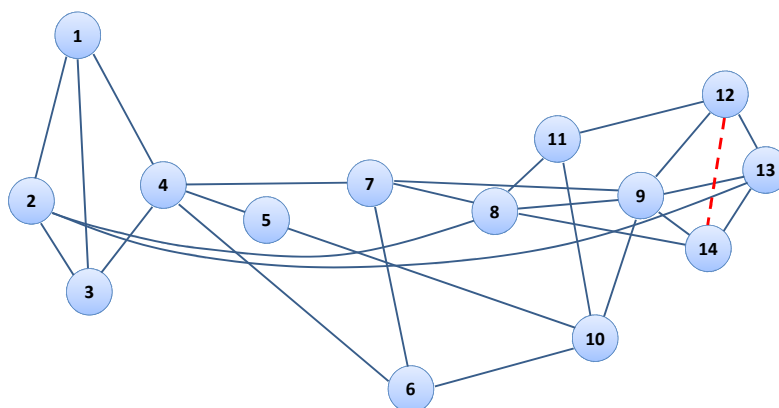
É importante ressaltar que uma vez que a RNA é treinada, uma avaliação de PB é determinística, mas os resultados de um simulador de eventos discretos são probabilísticos (pode variar a cada simulação). Neste exemplo, foi possível perceber que a proposição de novas métricas, como $I(\hat{\mathcal{F}})$ e CR , que são contribuições desta tese, são elementos-chave para aferir com precisão o desempenho de redes WDM que apresentam uma topologia física semelhante, visto que as outras métricas calculadas apresentam valores iguais ou muito próximos.

Foi adicionado um novo parâmetro de entrada sobre a RNA de 7 entradas de modo a permitir a estimativa de PB de redes com diferentes cargas de tráfego. O novo parâmetro é a carga a qual a rede está sujeita, em erlangs. A RNA foi treinada com dois tipos diferentes de carga: 100 e 200 erlangs. A nova RNA com 8 parâmetros de entrada apresentou $\overline{MSE} = 2,81 \cdot 10^{-4}$, $\sigma_{MSE} = 1,1 \cdot 10^{-5}$ e o tempo total para treinamento da rede foi em média 181 minutos. É desejável também neste estudo avaliar a capacidade de generalização da RNA, fornecendo cenários de carga de tráfego aos quais a rede não foi treinada.

Neste estudo foi usada a topologia original da rede NSFNet (vide Figure 3.13). A rede foi configurada para usar 20 e 40 comprimentos de onda e a carga foi variada de 80 até 220 erlangs. A Figura 4.10 exibe a PB fornecida pelo SIMTON e pela RNA. É possível observar que o método proposto neste trabalho pode estimar corretamente a PB de redes mesmo em cenários aos quais a RNA não foi apresentada, especialmente no caso de interpolação. O erro é maior para redes sujeitas a carga muito baixa. Felizmente, esta não é a região mais crítica



(a) Rede WDM com $W = 40$; $\delta = 38$ dB; $q = 0,31$; $CR = 0,42$; $CC = 0,24$; $\bar{c} = 83,37$; $I(\hat{\mathcal{F}}) = 0,91$.



(b) Rede WDM com $W = 40$; $\delta = 38$ dB; $q = 0,31$; $CR = 0,68$; $CC = 0,24$; $\bar{c} = 82,37$; $I(\hat{\mathcal{F}}) = 0,88$.

Figura 4.9: Comparação de duas redes WDM similares e o impacto das métricas na avaliação do desempenho.

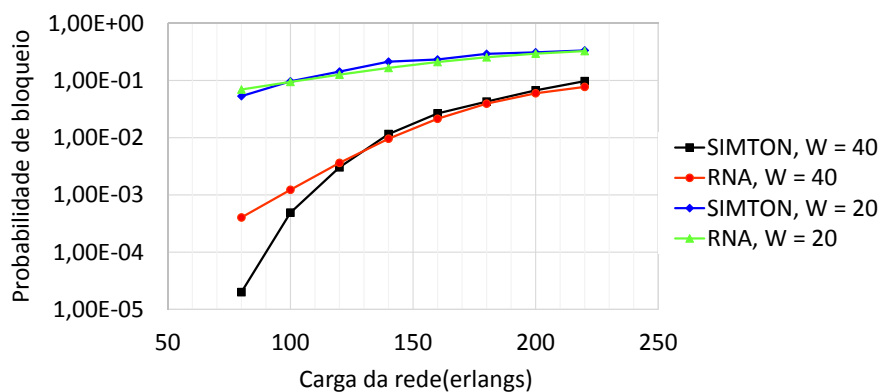


Figura 4.10: PB em função da carga da rede para o SIMTON e para a RNA treinada com 100 e 200 erlangs.

para redes de transporte e a nova proposta permite uma boa avaliação na região de interesse.

“SIMTON A” e “SIMTON B” apresentam um tempo médio de estimativa de PB de 3950 *ms* e 7553 *ms*, respectivamente. Por outro lado, uma vez que a RNA é treinada e as métricas topológicas são previamente calculadas, o tempo de execução para estimativa de PB usando a proposta deste trabalho é inferior a 1 *ms*. Ou seja, se forem considerados todos os efeitos da camada física no SIMTON e se a RNA é usada para avaliar uma rede óptica na qual se conhecem as características topológicas, o *speedup* obtido com a nova proposta é de até 7500 vezes para o caso NSFNet.

Discussão dos Resultados da Etapa Final do Estudo

A primeira análise de interesse neste grupo de resultados é a identificação do melhor conjunto para estimar PB, para ambas as variantes do problema, quando as 20 variáveis da Tabela 4.4 são consideradas. O método *forward-stepwise* foi aplicado para ambas as variantes do problema e para cada uma das três distribuições geográficas de redes. A Tabela 4.10 apresenta o resumo sobre as RNAs obtidas para cada caso. O detalhamento completo do processo de seleção de variáveis, incluindo outras combinações de variáveis, é apresentado no Apêndice A.

Tabela 4.10: Melhores modelos de regressão para estimativa de PB para redes FLTG, ARNES e MEDIANET.

Cenário	Tipo	Variáveis escolhidas	Média e desvio padrão de \overline{MSE}
FLTG	A	$W, \delta, q, \bar{c}_{km}, CR, L, \overline{\Delta OSNR}, \sigma_{\Delta OSNR}$	$5,43\text{E-}4 \pm 1,96\text{E-}5$
	B	$W, q, \bar{c}, I(\hat{\mathcal{F}}), L$	$5,30\text{E-}4 \pm 9,82\text{E-}6$
ARNES	A	$W, \delta, q, \bar{c}_{km}, CR, L, q_{km}$	$2,97\text{E-}4 \pm 8,65\text{E-}5$
	B	$W, q, \bar{c}, I(\hat{\mathcal{F}}), L$	$3,36\text{E-}4 \pm 4,31\text{E-}6$
MEDIANET	A	$W, \delta, q, \bar{c}_{km}, CR, L, \overline{\Delta OSNR}$	$3,13\text{E-}4 \pm 2,20\text{E-}6$
	B	W, CC, q, \bar{c}, L	$3,09\text{E-}4 \pm 4,30\text{E-}6$

Uma análise comparativa entre as linhas da Tabela 4.10 revela que a RNA possui erros de estimativa semelhantes para o problema Tipo A e para o problema Tipo B, mas a \overline{MSE} das redes FLTG é maior quando comparada aos demais casos. Este fato sugere que as redes com uma maior cobertura geográfica, e maior número de nós, são mais difíceis de serem representadas pela RNA. Contudo, em todos os cenários analisados, e para ambas as variantes do problema, foi obtido $\overline{MSE} < 10^{-3}$.

Quando são observadas as variáveis escolhidas para os três estudos de caso, percebe-se que os conjuntos de variáveis não são idênticos, mas algumas variáveis são recorrentes em todos os processos de seleção. Para o problema Tipo A, estão presentes em todos os cenários: número máximo de comprimentos de onda, fator de isolamento das portas do ROADM, densidade de enlaces, comprimento médio dos caminhos (em quilômetros), concentração de rotas e carga da rede. As demais variáveis que diferem de um estudo de caso para outro são usadas para fornecer precisão adicional ao estimador, considerando as particularidades de cada cenário.

No caso do problema Tipo B, as variáveis que estão presentes nos três cenários, são: o número máximo de comprimentos de onda, a densidade de enlaces, o comprimento médio dos caminhos (em saltos) e a carga da rede. Diversas outras combinações que apresentam um erro de estimativa próximos aos da Tabela 4.10, estão presentes no Apêndice A e revelam que outras variáveis podem ser usadas, conforme conveniência. Por exemplo, o uso de uma métrica que apresenta menor complexidade computacional pode ser adotada no lugar de outra que apresenta maior complexidade computacional, em detrimento de uma pequena variação nos erros de estimativa. É importante enfatizar que as métricas $I(\hat{\mathcal{F}})$ e CR, propostas no Capítulo 3, estão presentes em 5 das 6 melhores arquiteturas de RNAs obtidas na Tabela 4.10. Ou seja, as métricas propostas são elementos-chave para obtenção de estimativas precisas de PB por meio de uma RNA. Além disso, estas duas métricas aparecem em 27 das 49 combinações de variáveis apresentadas no Apêndice A, destacando que as métricas também são importantes para obtenção de RNAs mais simples, com menos variáveis e menor precisão.

Considerando as peculiaridades percebidas sobre os melhores estimadores para os casos individuais, um estudo adicional foi desenvolvido com o objetivo de investigar o desempenho de uma arquitetura de RNA única que considere a união dos melhores conjuntos de variáveis obtidos pelo casos particulares. O conjunto de variáveis independentes obtido possui 11 variáveis, a saber: número máximo de comprimentos de onda (W); fator de isolamento nos ROADMs (δ); carga de tráfego da rede (L); comprimento médio dos caminhos (\bar{c}), em quilômetros; concentração de rotas (CR); coeficiente de agrupamento (CC); entropia da DFT dos autovalores do Laplaciano ($I(\hat{\mathcal{F}})$); densidade de fibras ópticas, em quilômetros ($q(km)$); densidade de enlaces (q); margem média e desvio padrão de OSNR. A Tabela 4.11 apresenta os erros obtidos, para os três estudos de casos, quando esta RNA com 11 variáveis de entrada é considerada, e compara esta configuração com cada um dos três modelos específicos. Na Tabela 4.11, a coluna “Modelo específico” corresponde ao uso de uma camada de entrada específica, de acordo com a Tabela 4.10; a coluna “11 variáveis” corresponde ao uso de uma RNA contendo a união das variáveis obtidas nos estudos individuais. Uma comparação entre as duas colunas demonstra que as novas métricas consideradas proporcionam uma diminuição nos erros de estimativa em todos os cenários. Ou seja, é possível perceber que, embora o procedimento *forward-stepwise* tenha fornecido modelos específicos de apenas 7 ou 8 variáveis para cada cenário, o modelo de 11 variáveis oferece uma melhora acima de 5% em comparação com os modelos específicos. Isto é explicado pelo fato de que o processo de seleção considera a contribuição individual das variáveis e interrompe o processo quando $\Delta\overline{MSE} < 0,05$, mas quando são adicionadas duas ou mais variáveis em conjunto, a diminuição do erro não é afetada apenas pela contribuição individual das variáveis, mas também pela interação entre estas variáveis. Contudo, do ponto de vista prático, ambas as opções apresentadas na Tabela 4.11 podem ser usadas para estimar a PB das redes analisadas.

Uma análise interessante diz respeito ao custo computacional adicional requerido quando

Tabela 4.11: Comparação entre duas diferentes arquiteturas de RNA para estimativa de PB dos estudos de FLTG, ARNES e MEDIANET.

Estudo de caso	n	Média e desvio padrão de MSE	
		Modelo específico	11 variáveis
FLTG	41	$5,43\text{E-}4 \pm 1,96\text{E-}5$	$4,98\text{E-}4 \pm 2,63\text{E-}5$
ARNES	34	$2,97\text{E-}4 \pm 8,65\text{E-}5$	$2,90\text{E-}4 \pm 4,17\text{E-}5$
MEDIANET	18	$3,13\text{E-}4 \pm 2,20\text{E-}6$	$2,75\text{E-}4 \pm 1,01\text{E-}5$

a opção de 11 variáveis é usada no lugar das opções específicas de 7 e 8 variáveis. Para o caso MEDIANET, as variáveis adicionais são CC , \bar{c}_{km} , $I(\hat{\mathcal{F}})$, q_{km} e $\sigma_{\Delta OSNR}$. As variáveis adicionais sobre o caso MEDIANET possuem complexidade computacional $\mathcal{O}(n^2)$ (q_{km}), $\mathcal{O}(n^3)$ (CC , \bar{c}_{km} e $I(\hat{\mathcal{F}})$) e $\mathcal{O}(n^4)$ ($\sigma_{\Delta OSNR}$). Contudo, o fluxo mais custoso para a maioria destas métricas envolve passos já usados pelas outras 7 métricas, como encontrar os caminhos mínimos (já computada para $\overline{\Delta OSNR}$). De fato, o custo computacional acrescentado para o caso MEDIANET quando o modelo de 11 variáveis é usado em detrimento do modelo de 7 variáveis é devido ao cálculo dos autovalores para $I(\hat{\mathcal{F}})$. Mas como a RNA já precisa de um laço de $\mathcal{O}(n^4)$, o cálculo de $I(\hat{\mathcal{F}})$, que é de $\mathcal{O}(n^3)$, não oferece um aumento significativo no tempo de processamento. Uma análise semelhante pode ser feita para os casos ARNES e FLTG. Para efeito de padronização e para permitir uma comparação justa entre os diversos casos, os próximos estudos apresentados nesta subseção usam a arquitetura de 11 variáveis.

A Figura 4.11 apresenta um estudo sobre o impacto da variação da carga para a estimativa de PB considerando três diferentes abordagens para o problema Tipo A. As redes usadas na Figura 4.11 foram obtidas da base de dados de redes geradas para treinamento e foi escolhida a primeira rede que apresenta $1\text{E-}3 < \text{PB} < 1\text{E-}2$ para a menor carga. A rede óptica foi configurada para usar no máximo 20 ou 40 comprimentos de onda. Vale a pena mencionar que as RNAs foram treinadas apenas para três cargas diferentes. Portanto, os gráficos da Figura 4.11 também ilustram a capacidade das RNAs em interpolar os dados relacionados com cargas de rede não usadas no treinamento. Uma visão geral dos três cenários apresentados na Figura 4.11 sugere que a RNA acompanha a curva de PB *versus* carga original do SIMTON, mas apresenta dificuldade em estimar PB de redes que possuem muitos recursos e baixa ocupação (região de baixa PB para $W = 40$). Este comportamento pode ser entendido pelo fato de que, as redes nesta região de baixa carga e elevado número de comprimentos de onda, apresentam poucos bloqueios devido à indisponibilidade de comprimentos de onda e apresentam mais bloqueios relacionados com a degradação do sinal óptico.

A avaliação do melhor conjunto de variáveis é uma tarefa importante para obtenção do estimador baseado em RNA, mas uma preocupação adicional, ainda não explorada, é sobre o número mínimo de amostras necessário para obtenção de um estimador preciso. A Fig. 4.12 mostra o tempo médio de treinamento e a média dos erros quadráticos médios para diferentes números de redes consideradas na base de treinamento, para o problema Tipo A. A Fig. 4.13

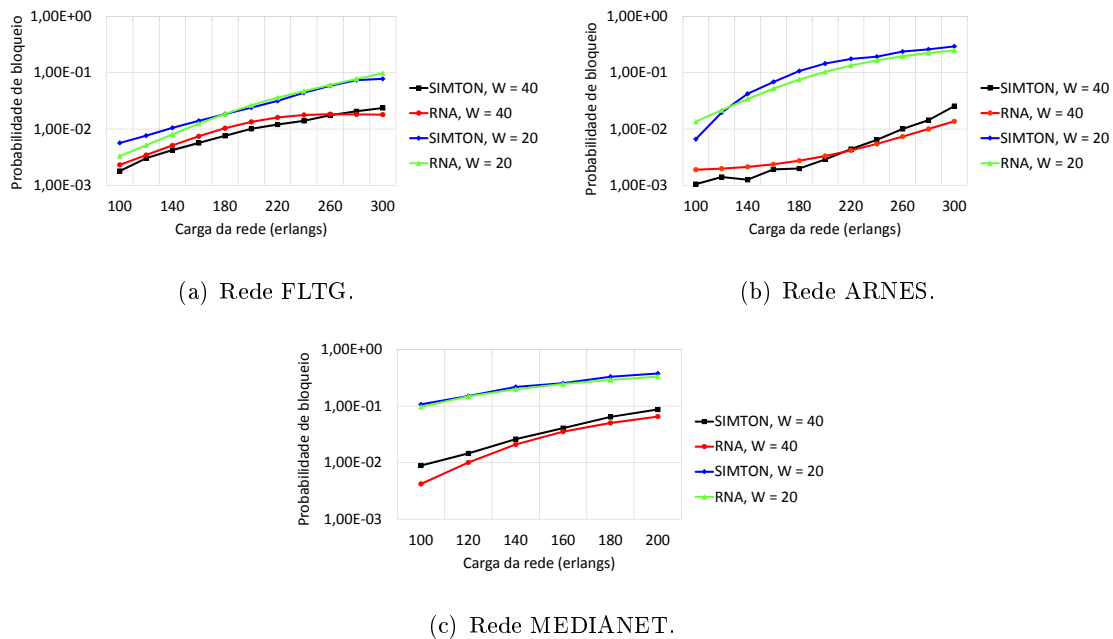
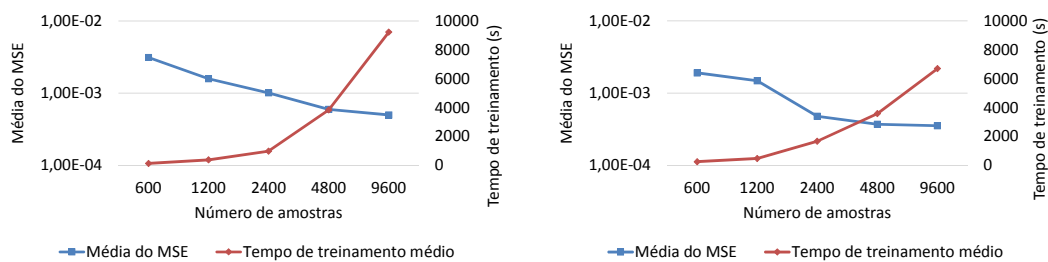


Figura 4.11: Estimativa de PB usando três diferentes abordagens e considerando variação da carga da rede, para o problema Tipo A.

apresenta a mesma informação para o problema Tipo B. O número de redes é sempre um múltiplo de 300 pois são consideradas 3 diferentes cargas de tráfego e 10 diferentes faixas de PB (vide Table 4.3).

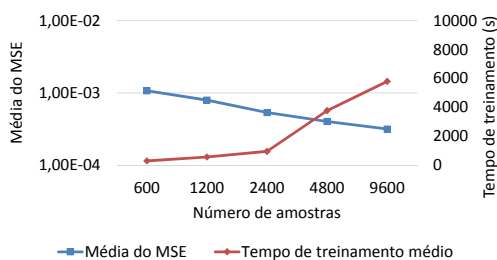
De acordo com a Fig. 4.12, o número mínimo de redes para treinar uma RNA para o problema Tipo A para obter $\overline{MSE} < 10^{-3}$ para FLTG, ARNES e MEDIANET é 4800, 2400 e 1200, respectivamente. De acordo com a Fig. 4.13, o número mínimo de redes para treinar uma RNA de modo a obter $\overline{MSE} < 10^{-3}$ para o problema Tipo B é 600 para os três cenários analisados. Uma comparação entre a Fig. 4.12 e a Fig. 4.13 revela que o problema Tipo A é mais difícil de ser representado por uma RNA quando comparado ao Tipo B, mesmo levando em consideração o número de variáveis adicionais usado para representar a camada óptica. Por exemplo, quando 1200 redes derivadas da FLTG são usadas para treinar uma RNA, é obtido $\overline{MSE} = 1,6E-3$ e $\overline{MSE} = 7,1E-4$ para o problema Tipo A e Tipo B, respectivamente. Por outro lado, para o mesmo exemplo da rede FLTG, é possível perceber que o tempo de treinamento é 31% menor para a abordagem usada no Tipo B quando comparada à abordagem usada para o Tipo A. Quando 9600 amostras são consideradas, o problema Tipo B exige 57% menos de tempo de treinamento.

Embora o tempo médio de treinamento da RNA seja uma preocupação importante quando um método alternativo de avaliação é proposto, uma maior atenção deve ser dada para o tempo de avaliação de uma rede individual. A RNA pode ser treinada por meio de um processo *offline*, então pequenas diferenças entre o tempo de treinamento não serão observadas por



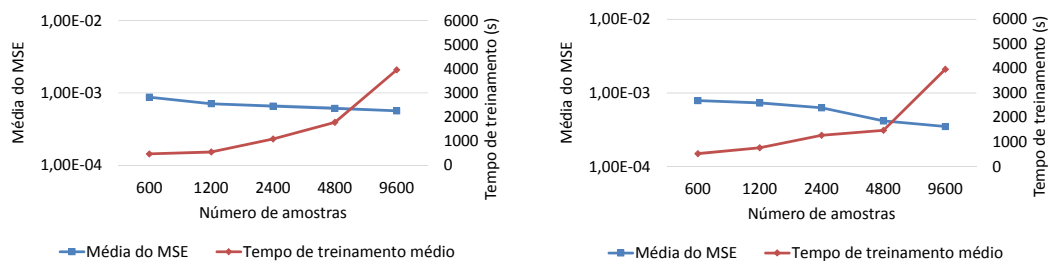
(a) Distribuição de nós da rede FLTG.

(b) Distribuição de nós da rede ARNES.



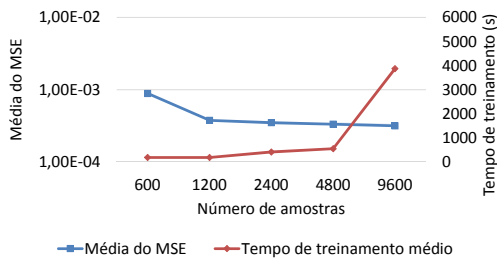
(c) Distribuição de nós da rede MEDIANET.

Figura 4.12: \overline{MSE} e tempo de treinamento para diferentes números de redes usadas no treinamento da RNA para o problema Tipo A.



(a) Distribuição de nós da rede FLTG.

(b) Distribuição de nós da rede ARNES.



(c) Distribuição de nós da rede MEDIANET.

Figura 4.13: \overline{MSE} e tempo de treinamento para diferentes números de redes usadas no treinamento da RNA para o problema Tipo B.

Tabela 4.12: Comparação entre o tempo médio de avaliação para simulações de rede e o tempo médio de avaliação para RNAs.

Rede	Tipo de Problema	Método de Estimativa	Tempo de Avaliação (ms)	Speedup
FTLG	A	Simulador de rede RNA	2203,04 130,10	17
	B	Simulador de rede RNA	1223,49 100,38	12
ARNES	A	Simulador de rede RNA	1634,76 58,33	28
	B	Simulador de rede RNA	868,92 41,57	18
MEDIANET	A	Simulador de rede RNA	480,00 6,68	72
	B	Simulador de rede RNA	274,41 4,28	64

um tomador de decisão quando uma ferramenta de projeto de redes é usada. Portanto, também foi avaliado o tempo médio de avaliação de redes após a fase de treinamento ser concluída. A Tabela 4.12 apresenta um resumo do tempo de avaliação para a abordagem tradicional, baseada em simulações de rede, e para a abordagem proposta nesta Tese, baseada em uma RNA. O tempo de avaliação da RNA apresentado na Tabela 4.12 considera que todas as métricas topológicas precisam ser calculadas no momento de avaliação, ou seja, é o pior caso para a abordagem baseada em RNA. Por outro lado, a avaliação no SIMTON usa 10^7 chamadas para os três cenários, com parada após 1000 bloqueios, e não considera efeitos não-lineares (na Seção 4.2.3 é considerado o efeito FWM). De acordo com a Tabela 4.12 a abordagem proposta oferece expressivos *speedups* para todos os cenários analisados. Além disso, para o problema Tipo A, a melhoria no tempo de resposta é ainda mais expressiva quando comparada ao problema Tipo B. Em outras palavras, quando uma avaliação de uma rede óptica exige a consideração das penalidades da camada física, um método de avaliação alternativo, baseado em RNA é ainda mais vantajoso quando comparado com as simulações de rede.

Quando a RNA estima uma PB, são usadas medidas topológicas da rede como parâmetros de entrada, mas alguns parâmetros de simulação da camada óptica não são considerados no treinamento (vide Tabela 4.2). Portanto, a RNA fornece uma estimativa de PB precisa para uma rede WDM com os parâmetros de rede que foram usados nos experimentos. Se for necessário avaliar uma rede óptica em um cenário completamente diferente, um novo processo de treinamento pode ser necessário. Contudo, para propósitos práticos relacionados com análise de redes WDM, é muito comum permanecer com alguns parâmetros da rede imutáveis por um longo período e realizar várias simulações com estes parâmetros predeterminados para aferição de QoT para diferentes configurações de rede. Contudo, é interessante conhecer em quais situações é mais apropriado usar a proposta deste trabalho, mesmo considerando a

necessidade de um treinamento inicial da rede, ou seja, considerando o pior caso. Se um conjunto de s simulações de redes WDM são necessárias, deseja-se encontrar um valor mínimo de $s = s_{min}$ simulações no qual é melhor usar a proposta deste trabalho. O tempo total necessário para avaliação de s redes WDM usando a RNA é de $\Delta t^{RNA} = \Delta t_{training}^{RNA} + s \cdot \Delta t_{evaluation}^{RNA}$. O tempo necessário para avaliar s redes WDM usando um simulador similar ao SIMTON é de $\Delta t^{SIMTON} = s \cdot \Delta t^{SIMTON_i}$. De modo a obter um ganho líquido no uso da nova proposta, mesmo considerando o pior caso, é necessário que $\Delta t^{RNA} < \Delta t^{SIMTON}$. Se for considerado o tempo de treinamento para o problema Tipo A com 1200 amostras é possível calcular s_{min} para o caso MEDIANET por meio de

$$\Delta t_{training}^{RNA} + s_{min} \cdot \Delta t_{evaluation}^{RNA} < s_{min} \cdot \Delta t^{SIMTON_i}, \quad (4.6)$$

$$572 \cdot 1.000 + 6,68 \cdot s_{min} < 480,00 \cdot s_{min} \Rightarrow s_{min} > 1.208.$$

Portanto, se for considerado o método proposto nesta Tese, no pior caso após 1.209 simulações já é vantajoso usar a nova metodologia. O mesmo procedimento aplicado aos cenários ARNES e FLTG (usando 4800 amostras) fornece $s_{min} > 2270$ e $s_{min} > 1.849$, respectivamente. Em geral, centenas de milhares de avaliações podem ser necessárias em uma ferramenta de planejamento de redes baseada em simulações tipo Monte Carlo. Um dos resultados contidos em [22], destaca que 450.000 avaliações de redes WDM podem ser necessárias para uma completa convergência em uma ferramenta de planejamento de rede. Outros estudos, por questões de limitação de custo computacional, consideraram 50.000 diferentes configurações de rede [99]. Portanto, a exigência de 1209-2270 avaliações é muito vantajosa quando são considerados os trabalhos anteriores sobre planejamentos por processos iterativos. Além disso, o valor obtido para s_{min} é aceitável se for considerado que um novo processo de treinamento ocorrerá em situações raras. Como exemplos de tais situações, podem ser citadas o desenvolvimento de um novo dispositivo óptico ou de um novo modelo para a camada física.

Uma última análise sobre o desempenho de diferentes abordagens para estimativa de PB é ilustrada na Figura 4.14. Os gráficos de *box-plots* apresentados na Figura 4.14 foram criados a partir de 20 rodadas de avaliações independentes de 600 diferentes redes ópticas. “SIMTON A” é o método de estimativa de PB baseado em simulações de Monte Carlo limitadas a 10^7 requisições de chamadas ou até que 1000 bloqueios ocorram. “SIMTON B” é baseado em simulações de Monte Carlo limitadas a 10^5 requisições de chamadas ou até que 500 bloqueios ocorram. A arquitetura da RNA usada para o comparativo da Figura 4.14 foi obtida por meio do treinamento de 4800 redes e possui 11 variáveis na camada de entrada. A Figura 4.14 confirma a análise já realizada para a rede NSFNet (Figura 4.8), ou seja, a abordagem baseada em RNA é uma alternativa interessante, pois apresenta erros de estimativa inferiores aos obtidos por simulações de redes que usam poucas requisições de chamadas. Para todos os cenários de

rede analisados, a abordagem baseada em RNA apresenta \overline{MSE} entre a configuração “SIMTON A” e “SIMTON B”. Considerando que o tempo necessário para avaliação pela RNA é inferior aos métodos baseados em simulação, é possível estabelecer um compromisso entre precisão e custo computacional para a RNA e a opção “SIMTON A”, mas quando “SIMTON B” e o método baseado em RNA são comparados, este último é mais vantajoso. Além disso, o uso de poucas chamadas em simulações para estimativa de PB de redes que apresentam uma grande quantidade de nós mostrou-se inadequado, visto que $1E-2 < \overline{MSE} < 1E-3$ para os casos ARNES e FLTG.

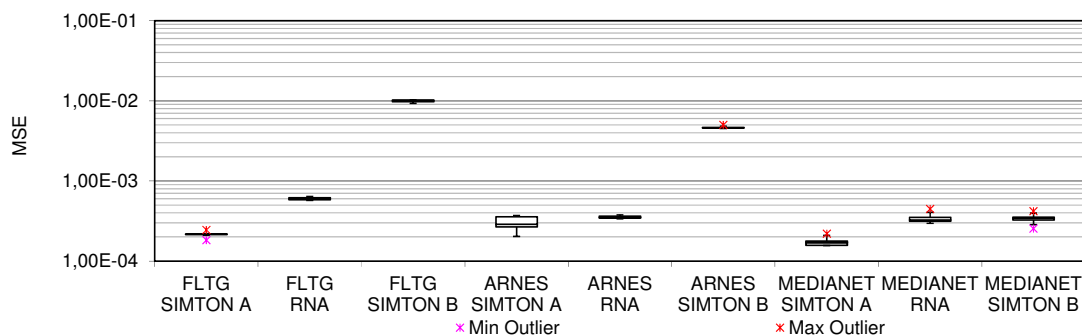


Figura 4.14: *Box-plots* da média do erro médio quadrático, considerando 20 avaliações para cada método de estimativa de PB, para redes ARNES, FLTG e MEDIANET.

4.2.4 Considerações Finais sobre Estimativa de PB Usando RNA

Diversos métodos para estimativa de probabilidade de bloqueio de redes ópticas já foram propostos na literatura. Contudo, a escolha por um método de estimativa em particular deve levar em consideração o compromisso entre precisão da estimativa (erro de estimativa) e custo computacional associado (tempo gasto no processo de estimação).

Esta Tese propõe o uso de RNA para estimar PB de redes ópticas. Os resultados demonstram que a abordagem proposta oferece um bom compromisso em termos de precisão e custo computacional quando comparado com abordagens disponíveis na literatura. A nova abordagem oferece um tempo de execução até 72 vezes menor quando comparado com abordagens baseadas em simulação de rede. Os erros de estimativa obtidos pela nova abordagem possuem a mesma ordem de grandeza dos erros associados com simulações de rede com grande número de chamadas ($\overline{MSE} \approx 10^{-4}$). Dependendo do número de chamadas usadas nas simulações, a precisão da RNA pode ser considerada ligeiramente inferior ou ligeiramente superior à precisão fornecida pelas simulações. A principal motivação para o desenvolvimento de uma nova abordagem para estimativa de PB foi o uso do novo método em processos iterativos de planejamento de redes, nos quais são necessárias milhares de estimativas de desempenho de

redes. O método de estimativa proposto neste capítulo será usado no Capítulo 5 como parte integrante do processo de planejamento de redes ópticas.

Um primeiro estudo sobre a viabilidade de usar RNA para estimativa de PB de redes foi apresentado em 2013, no “*International Microwave and Optoelectronics Conference*”, e um artigo foi publicado nos anais deste evento [92]. A proposta de uma metodologia para seleção de variáveis da camada de entrada da RNA e um estudo sobre a influência do tráfego para estimativa de PB foi publicado no “*Journal of Optical Communications and Networking*” [85], em 2015. Um estudo sobre a proposição de novas métricas para consideração de efeitos da camada física e estudos de casos em redes implantadas foi apresentado no “*17th International Conference on Transparent Optical Networks ICTON*”, e publicado nos anais deste mesmo evento [100]. Um estudo sobre o impacto do número de redes usadas para treinamento da RNA e a proposição de uma arquitetura para estimativa de PB em cenários que não consideram efeitos da camada física foi aceito para a edição de 2015 do evento “*SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference*” [101].

Capítulo 5

Planejamento de Redes Ópticas

O projeto da topologia de uma rede compreende a determinação de quais nós estarão conectados. Dependendo das premissas e das características da rede, o projeto da topologia pode ser considerado uma tarefa extremamente difícil, e, por este motivo, este problema vem sendo exaustivamente estudado nos últimos anos. Os primeiros esforços para resolver este problema consideram a determinação de árvores geradoras mínimas (*Minimum Spanning Trees* - MST). O projeto de topologia de uma rede é tipicamente associado com problemas de MST com restrições, como por exemplo, o requisito de que cada nó esteja conectado a pelo menos outros dois nós. Restrições desta natureza favorecem a adoção de estratégias de sobrevivência, como proteção e restauração, mas levam o projeto de topologia para a categoria de um problema NP-difícil [102].

O planejamento completo de uma rede óptica implica na escolha da conexão entre os nós da rede por meio de fibra óptica e, adicionalmente, na escolha de dispositivos ópticos. Esta escolha deve considerar que um elevado número de conexões irá aumentar o custo de implantação da rede (*Capital Expenditure* - CAPEX), mas poderá também suportar uma carga de tráfego maior se as conexões forem definidas apropriadamente. Em uma rede transparente, o sinal permanece no domínio óptico durante o caminho inteiro entre a origem e o destino da transmissão, sem que nenhuma conversão óptica-elétrica-óptica (O-E-O) ocorra. Apesar de este tipo de rede apresentar um menor custo quando comparada com redes opacas e translúcidas, a acumulação de penalidades no sinal em caminhos ópticos mais longos pode diminuir a Qualidade de Transmissão (QoT) devido aos severos efeitos da camada óptica. Em geral, a degradação do sinal também está relacionada com a qualidade dos dispositivos empregados e consequentemente, implicam no custo da rede. De fato, o planejamento de uma rede óptica deve considerar o compromisso entre dispositivos com custo mais baixo que irão conferir um desempenho inferior à rede e entre dispositivos mais caros que podem fornecer um desempenho global superior à rede óptica. Portanto, o planejamento completo de uma rede óptica, incluindo o projeto da topologia física e especificação de dispositivos é um problema

de otimização multiobjetivo típico.

Algoritmos evolucionários têm sido usados com sucesso para a solução de problemas de otimização com restrições e muitos paradigmas já foram propostos, dentre os quais podem ser destacados: Algoritmos Genéticos, Evolução Diferencial, Estratégias de Evolução e Programação Genética [103]. Por outro lado, problemas do mundo real apresentam desafios adicionais que precisam ser tratados pelos algoritmos de otimização. Primeiramente, problemas do mundo real geralmente apresentam muitos objetivos conflitantes que precisam ser otimizados simultaneamente. Segundo, a tarefa de avaliar uma solução candidata geralmente é computacionalmente custosa em termos de tempo de execução para problemas reais. Já foi demonstrado para diversos casos que algoritmos evolucionários também podem ser aplicados a problemas com dois ou mais objetivos conflitantes [104]. Os Algoritmos Evolucionários Multi-objetivos (AEMO) geram um conjunto de soluções e a maioria destes algoritmos usam o conceito de dominância. Alguns AEMOs também já foram aplicados para a solução do problema Árvore Geradora Mínima Multi-objetivo (MO-MST) [105, 106] e em planejamento de redes ópticas [49, 22]. Alguns trabalhos recentes sobre otimização multiobjetiva de problemas reais têm focado no estudo de modelos substitutos (*Surrogate Models*) para avaliação de soluções de modo a diminuir o tempo total de execução dos algoritmos evolucionários [107, 108].

Conforme apresentado no Capítulo 3, os avanços em Ciência das Redes permitiram o desenvolvimento de novas soluções para problemas do mundo real que podem ser modelados como grafos. Uma das hipóteses desta Tese é de que as métricas topológicas desenvolvidas recentemente e os modelos de geração de redes podem ser usados com sucesso para construção de ferramentas de planejamento de redes ópticas, especialmente no que concerne ao projeto de topologia física.

Neste capítulo será proposta uma metodologia evolucionária multiobjetiva para tratar o problema de projeto de topologia física de redes ópticas. Nesta metodologia são usados conceitos de Ciência das Redes e técnicas de aprendizagem de máquinas. Este capítulo está organizado da seguinte forma: a Seção 5.1 apresenta uma fundamentação teórica sobre otimização multiobjetiva; a Seção 5.2 apresenta uma descrição formal do problema que é tratado neste capítulo; a Seção 5.3 descreve a proposta desta Tese para o planejamento de redes ópticas.

5.1 Otimização Multiobjetiva

Esta seção apresenta conceitos gerais de otimização multiobjetiva e sobre o algoritmo evolucionário que foi usado como base para o desenvolvimento na proposta de planejamento desta Tese. A subseção 5.1.1 apresenta os conceitos e a terminologia básica que é usada em otimização multiobjetiva. A subseção 5.1.2 apresenta as métricas usadas para avaliação de qualidade de soluções fornecidas por algoritmos de otimização multiobjetivos. A subseção 5.1.3 apre-

senta o algoritmo NSGAI, que é usado como base para a metodologia evolucionária que é proposta pela Tese.

5.1.1 Conceitos Básicos

A otimização de problemas com objetivo único está relacionada com encontrar valores para um vetor de variáveis de decisão $\hat{x} = \{x_1, \dots, x_n\}$ em um espaço de decisão Ω que minimiza (ou maximiza) o valor de uma variável independente ou função objetivo y ou $f(\hat{x})$. Neste tipo de problema, a noção de solução “ótima” é clara pois existe um \hat{x} que apresenta o melhor valor para y .

Contudo, diversos problemas do mundo real apresentam objetos conflitantes que precisam ser considerados. Por exemplo, o atraso de uma rede de computadores pode ser diminuído por meio da atualização dos dispositivos usados na rede. Contudo, uma troca de equipamentos representa um custo adicional para o projeto da rede. Neste problema, o vetor de variáveis de decisão poderia ser os dispositivos da rede e os dois objetivos conflitantes seriam o custo total para atualização da rede e o atraso da rede. Na Figura 5.1, as soluções A, B, C, D e E representam escolhas diferentes em atualização de dispositivos e são preferíveis às soluções F e G pois o primeiro grupo oferece vantagens em um objetivo ou em ambos. Contudo, se as soluções $A - E$ são comparadas entre si, não é possível decidir qual solução é melhor. Por exemplo, a solução E é mais barata do que a solução A , mas o atraso da rede da solução A é menor. Na terminologia de Pareto [104], as soluções $A - E$ **dominam** as soluções F e G ou $\{A, B, C, D, E\} \preceq \{F, G\}$. Além disso, as soluções $A - E$ são incomparáveis entre si e elas estão no mesmo conjunto de soluções não-dominadas que é conhecido como uma **frente de Pareto**.

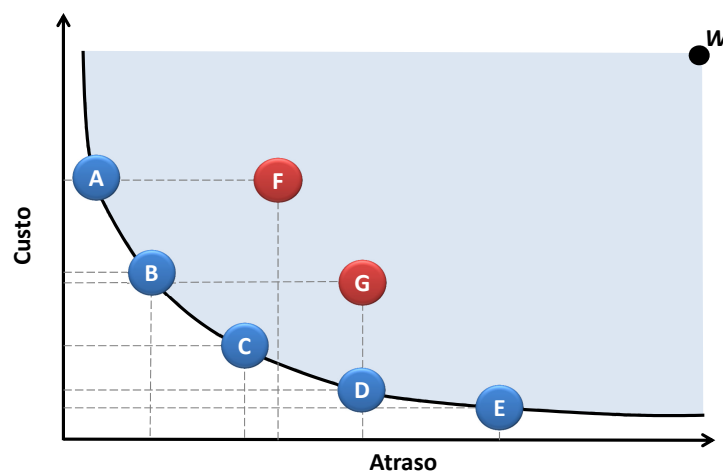


Figura 5.1: Um problema com duas funções objetivo: custo e atraso da rede. A frente de Pareto ou superfície de compromisso está destacada pela linha que une os pontos $A - E$.

Quando um problema possui duas ou mais funções objetivos, a noção de otimalidade muda, pois a meta é encontrar a melhor relação de compromisso ao invés de uma solução única, como ocorre em otimização com objetivo único. Algoritmos de otimização multiobjetivos foram propostos para encontrar a melhor frente de Pareto possível para diversos problemas. Portanto, a meta para problemas com dois ou mais objetivos conflitantes é encontrar um conjunto Pareto ótimo \mathcal{P}^* , tal que

$$\mathcal{P}^* = \{\hat{x} \in \Omega \mid \nexists \hat{x}' \in \Omega, F(\hat{x}') \preceq F(\hat{x})\}. \quad (5.1)$$

5.1.2 Métricas para Avaliação de Desempenho de AEMO

Diversas métricas já foram propostas para avaliação da qualidade de soluções oferecidas pelos algoritmos de otimização multiobjetivos [109, 110, 111, 112]. A escolha de métricas para uso em um estudo de otimização multiobjetiva deve considerar a inclusão de pelo menos duas métricas: uma para avaliação de convergência das soluções e outra para avaliação da diversidade de soluções. Além disso, algumas métricas propostas na literatura, como *generational distance* [110], são aplicáveis apenas a problemas nos quais o Pareto ótimo é conhecido, o que não é verdade para o problema considerado nesta Tese. Portanto, neste trabalho serão usadas três métricas da literatura que combinam avaliação de convergência e diversidade das soluções e não exigem o conhecimento prévio do Pareto ótimo. Será usada a métrica *Coverage set* para analisar a capacidade de convergência dos algoritmos. A métrica *Maximum Spread* será usada para avaliar o espalhamento de soluções na frente de Pareto. Também será usado um indicador híbrido, denominado *Hypervolume*, para medir simultaneamente a capacidade de convergência e diversidade das soluções. Segue uma breve descrição destas métricas:

1. Comparação de cobertura de dois conjuntos (*Coverage set - CS*):

A comparação de cobertura de dois conjuntos [113] consiste no mapeamento de dois conjuntos (X_1, X_2) para um número no intervalo $[0, 1]$ de acordo com

$$CS(X_1, X_2) = \frac{|x_2 \in X_2; \exists x_1 \in X_1 : x_1 \preceq x_2|}{|X_2|}. \quad (5.2)$$

Se todos os pontos em X_1 dominam todos os pontos em X_2 então $CS = 1$. $CS = 0$ implica o contrário. CS para o exemplo ilustrado na Figura 5.1 resulta em $CS(\{A, \dots, E\}, \{F, G\}) = 1$ e $CS(\{F, G\}, \{A, \dots, E\}) = 0$.

2. *Hypervolume - HV*:

O *Hypervolume* foi proposto por Zitzler e Thiele [113] em 1999. Esta métrica é usada para avaliar a área destacada na Figura 5.1. O ponto W , conhecido como *datum point*, é usado como uma referência para avaliação da métrica. É possível perceber que frentes

de Pareto com soluções espalhadas e que apresentaram boa convergência irão conferir valores mais elevados para a métrica *Hypervolume*.

3. *Maximum Spread - MS*:

A métrica *Maximum Spread (MS)* foi proposta por Zitzler *et al.* [113] e estima a extensão da frente de Pareto. Para problemas com dois objetivos, *MS* coincide com a distância Euclidiana entre as duas soluções extremas da frente de Pareto. *MS* para o exemplo ilustrado na Figura 5.1 é a distância Euclidiana entre as soluções *A* e *E*.

5.1.3 O Algoritmo NSGAI

Em 2002, Deb *et al.* [114] propuseram o *Non-dominated Sorting Genetic Algorithm II* (NSGAI). Desde 2002, diversos estudos em problemas do mundo real demonstraram que NSGAI é uma meta-heurística adequada para resolver problemas multiobjetivos combinatoriais, especialmente se dois objetivos são considerados [104]. Para problemas com mais de quatro objetivos conflitantes, conhecidos na literatura inglesa como *many objectives*, Deb *et al.* propuseram recentemente uma versão melhorada do NSGAI, que ficou conhecida como NSGAIII [115]. Considerando que a modelagem usada para o problema estudado nesta Tese possui dois objetivos conflitantes, foi adotado o algoritmo NSGAI como meta-heurística básica para as propostas que serão apresentadas. A escolha pelo algoritmo NSGAI foi fundamentada em estudos anteriores sobre projeto de redes ópticas. O estudo realizado por Araujo *et al.* [99] comparou o uso de diversas meta-heurísticas, incluindo NSGAI, SPEA2, PESAII, MODE, PAES, e mostrou que os algoritmos NSGAI e SPEA2 oferecem os melhores resultados e entre estas duas últimas não há melhoria perceptível. Além disso, outro estudo sobre projeto de redes ópticas aplicou o algoritmo NSGAI e demonstrou que a melhoria na qualidade das soluções ocorre primariamente devido às adaptações realizadas nos operadores genéticos e em menor intensidade devido à meta-heurística adotada [22].

O pseudocódigo do NSGAI é mostrado no Algoritmo 4. Para um melhor entendimento do procedimento, considere a execução do NSGAI com um tamanho de população $N = 7$ e com as soluções que estão apresentadas na Figura 5.2(a). Na Figura 5.2(a), as soluções claras correspondem às soluções da geração anterior e as soluções escuras correspondem às soluções encontradas na geração atual (descendentes). Portanto, o NSGAI usa o conjunto formado pelos indivíduos da geração anterior e pelos descendentes para obtenção de um conjunto de tamanho $2N$. Este conjunto é classificado em classes ou frentes de Pareto, atribuindo um valor de aptidão 1 aos indivíduos não dominados, 2 às soluções da próxima frente de Pareto e assim por diante. Desta forma, a classificação consiste em obter os indivíduos com os menores valores. Este passo está ilustrado pelas

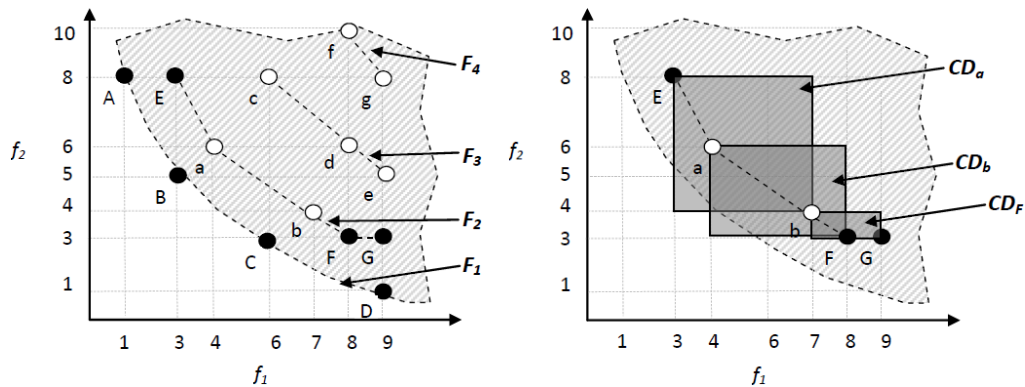
indicações $\{F_1, \dots, F_4\}$ na Figura 5.2(a). Após essa etapa de classificação, o conjunto é truncado para permanecer com o tamanho de população original ($N = 7$). Nos casos em que o truncamento precise descartar indivíduos de uma mesma frente de Pareto, é usado como critério de descarte uma estimativa de densidade utilizando o método *CD - Crowding Distance* ou “distância da multidão”. A medida *CD* de uma solução corresponde ao tamanho de um cubóide formado pelas soluções adjacentes à solução considerada. Um valor elevado de *CD* indica que a solução está contida em uma região pouco densa, e em contrapartida, um valor baixo desta métrica corresponde a uma solução contida em uma região mais densa. No exemplo apresentado na Figura 5.2, foram encontradas 9 soluções somando-se as frentes F_1 e F_2 . Portanto, é necessário descartar duas soluções da frente F_2 . Com a finalidade de manter um espalhamento máximo em cada frente de Pareto, o método *CD* atribui às soluções mais afastadas o valor máximo da métrica de densidade, ou seja, as soluções E e G da Figura 5.2 são automaticamente levadas para a próxima geração. Para completar o processo de seleção, é necessário decidir entre as soluções a , b e F . Formando um cubóide para cada solução, percebe-se que a solução a apresentou maior valor de *CD*, conforme ilustrado na Figura 5.2(c). Portanto, a próxima geração será formada pelas soluções $\{A, B, C, D, E, G, a\}$.

Algoritmo 4 O Algoritmo NSGAI.

```

1: procedimento NSGAI
2:   Crie uma população inicial POP
3:   para  $i \leftarrow 0$ ;  $i < I_{max}$ ;  $i++$  faça
4:     para  $s \leftarrow 0$ ;  $s < P$ ;  $s \leftarrow s + 2$  faça
5:       Selecione dois pais,  $P_1$  e  $P_2$ 
6:       Execute cruzamento entre  $P_1, P_2$  e gere os descendentes  $D_1$  e  $D_2$ 
7:       Faça mutação com  $D_1$  e  $D_2$ 
8:       Avalie  $D_1$  e  $D_2$  e inclua no conjunto de descendentes DESC
9:     fim para
10:    Crie um conjunto UNIAO a partir de POP e DESC
11:    Separe as frentes de Pareto usando o conceito de dominância
12:    Crie um conjunto vazio POP
13:    enquanto o tamanho de POP  $< P$  faça
14:      Ordene as soluções da frente de Pareto por CD
15:      para cada solução no conjunto ordenado faça
16:        se comprimento de POP  $< P$  então
17:          Inclua a solução no conjunto POP
18:        fim se
19:      fim para
20:      Analise a próxima frente de Pareto
21:    fim enquanto
22:  fim para
23: fim procedimento

```



(a) Classificação de soluções não-dominadas utilizada pelo NSGAI. (b) Avaliação de *crowding distance* de soluções que estão na mesma frente de Pareto.

Figura 5.2: Ilustração do mecanismo de seleção utilizado pelo NSGAI.

A Figura 5.3 resume o esquema de seleção de sobreviventes do NSGAI. Após unir as soluções da população anterior (P_t) e a população atual (D_t), as soluções são classificadas usando dominância, obtendo as frentes de Pareto F_1, F_2, F_3 e F_4 . As soluções de F_1 são automaticamente levadas para a próxima geração (P_{t+1}). As soluções de F_2 com maior valor de CD também são levadas para a próxima geração, até completar a população P_{t+1} com o tamanho $|P_{t+1}| = N = 7$. Portanto, as soluções descartadas são as soluções de F_3, F_4 , e as soluções de F_2 com baixo valor de CD . O comportamento observado durante as execuções do NSGAI é que nos primeiros passos é muito usada a classificação por dominância e após várias gerações, muitas soluções são geradas na mesma frente de Pareto, de modo que o critério para seleção das soluções passa a ser a distância da multidão.

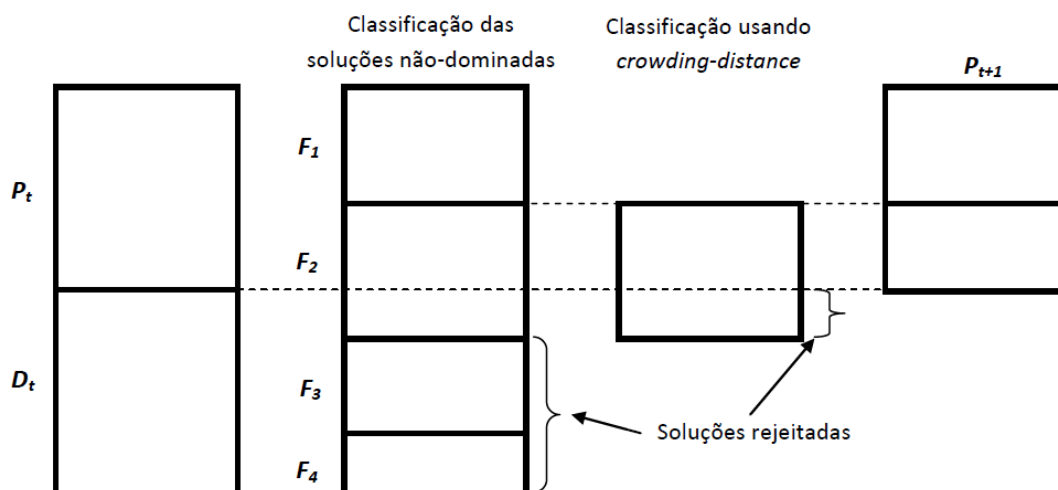


Figura 5.3: Modelo para seleção de sobreviventes usado no NSGAI.

5.1.4 Otimização Assistida por Modelos Substitutos

A computação evolucionária assistida por modelos substitutos faz uso de modelos eficientes e aproximados, frequentemente denominados na literatura de meta-modelos ou *surrogates*. Os meta-modelos são usados para diminuir o tempo de processamento para problemas de otimização computacionalmente custosos. Trabalhos recentes demonstram que meta-modelos podem ser usados com sucesso para resolver problemas do mundo real, como a otimização de projetos aerodinâmicos [116] e projetos de medicamentos [117].

Estudos nesta área tornaram-se populares após 1990 [118, 119]. Diversas estratégias e modelos computacionais foram desenvolvidos, incluindo métodos polinomiais, processos Gaussianos e redes neurais [120]. Meta-modelos podem ser usados em conjunto com a função de aptidão original para prevenir que o algoritmo evolucionário seja enganado por um falso ótimo introduzido quando apenas o meta-modelo é usado. Além disso, múltiplos meta-modelos podem ser combinados. Por exemplo, a um conjunto de redes neurais foi usado por Jin e Sendhoff [121], e Tenne e Armfield [122] propuseram o uso de polinômios, *support vector machines* (SVMs) e redes neurais de forma combinada.

Outra importante preocupação na área é o gerenciamento dos meta-modelos. As técnicas de gerenciamento de meta-modelos para avaliação de aptidão de soluções podem ser divididas em: técnicas baseada em indivíduos, técnicas baseadas em geração e técnicas baseadas em população [120]. Nas técnicas baseadas em geração, os meta-modelos são usados em algumas gerações apenas e nas demais gerações é usada a função de aptidão original. Na abordagem baseada em indivíduo, os meta-modelos são usados para alguns indivíduos e a função de aptidão original é usada para os demais. No esquema baseado em população, mais de uma subpopulação coevolui. Outra questão essencial é quais indivíduos devem ser avaliados ou reavaliados usando a função de aptidão original e algumas possibilidades já foram analisadas: é possível focar em indivíduos que possuam elevados valores de aptidão e elevada precisão na estimativa de modo a aumentar a frequência de uso dos meta-modelos [123]; outra abordagem é realizar um agrupamento da população de acordo com características e re-avaliar os indivíduos que estejam no centro do agrupamento [121]; é possível selecionar indivíduos que possuem um elevado grau de incerteza na aproximação da função para serem re-avaliados [124].

Finalmente, embora a precisão da estimativa não seja o único critério a considerar quando meta-modelos são usados, é desejável aumentar a qualidade destes modelos substitutos sempre que possível. Alguns trabalhos relatam avanços neste sentido com o foco em: regularização do modelo de rede neural [123]; otimização da estrutura e dos parâmetros do meta-modelo [125]; construção de um meta-modelo em um novo espaço de decisão com menos variáveis [126]. Uma revisão completa da literatura sobre computação evolucionária assistida por meta-modelos pode ser obtida em [120].

Diferentes formas de avaliação para enlaces ópticos e redes ópticas foram apresentadas no Capítulo 2. A Figura 5.4 apresenta alguns destes métodos ilustrados em termos de compro-

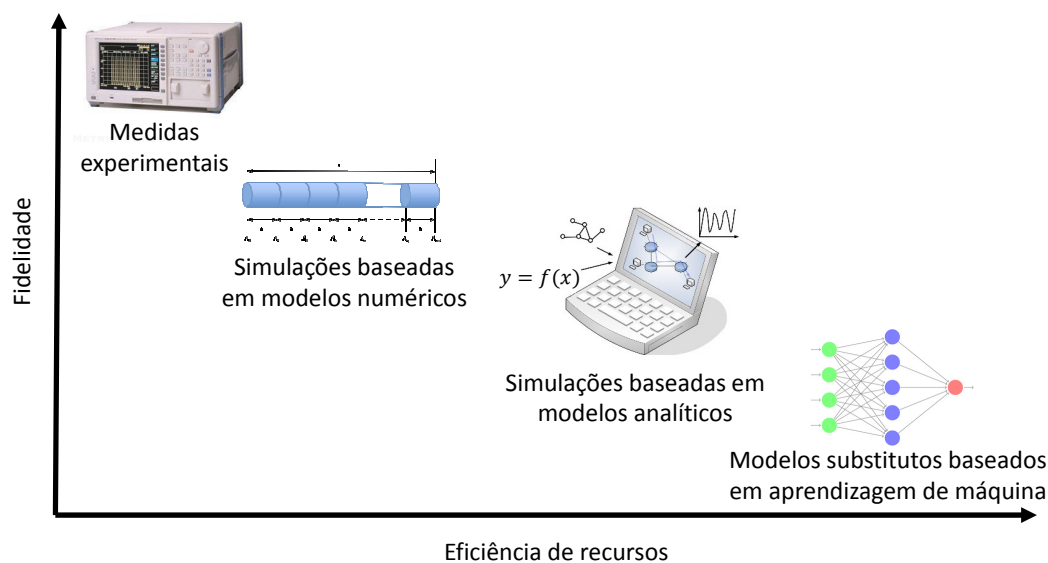


Figura 5.4: Exemplo de compromisso entre fidelidade das estimativas e eficiência de recursos para modelos substitutos em avaliação de redes ópticas.

misso entre eficiência de recursos e fidelidade da avaliação com relação à realidade. As medidas experimentais apresentam grande fidelidade com relação à qualidade do sinal de um enlace óptico, mas o seu uso prático envolve a implantação de uma infra-estrutura na qual a medida será obtida e o custo financeiro desta operação é impraticável se o objetivo final é a avaliação de uma infra-estrutura candidata. Portanto, a estratégia mais comum para avaliação de uma rede em um processo embrionário de planejamento é o uso de simulações computacionais. As simulações, por sua vez, apresentam diferentes compromissos entre tempo de processamento e fidelidade do modelo. As simulações baseadas em modelos numéricos são mais precisas, mas exigem maior tempo de processamento. Por outro lado, as simulações baseadas em modelos analíticos simplificados são menos precisas, mas exigem menor tempo de processamento. A proposta desta Tese para estimativa de PB de redes ópticas é baseada no uso de aprendizagem de máquina para construção de modelos alternativos que ofereçam menor tempo de processamento, mas mantendo a fidelidade próxima à apresentada por simulações baseadas em modelos analíticos.

5.2 O Problema de Planejamento de Redes Ópticas

O problema que é tratado neste capítulo pode ser enunciado da seguinte forma: dada a posição dos nós, uma matriz de tráfego (T) e um algoritmo de RWA; a meta é determinar a topologia física e a especificação dos equipamentos ópticos que devem ser usados de modo a

minimizar simultaneamente o custo de implantação da rede (CAPEX - *Capital Expenditure*) e a probabilidade de bloqueio da rede, sujeito aos diversos critérios de QoT considerados para o estabelecimento de caminhos ópticos. Os seguintes parâmetros da rede são usados como variáveis de decisão: os possíveis enlaces da topologia física, o tipo de ROADM usado na rede e o número máximo de comprimentos de onda disponíveis por enlace.

A posição dos nós é definida por $P = (x_i, y_i)$, em que $(i \in 1, 2, \dots, n)$, x_i e y_i são as coordenadas cartesianas do i -ésimo nó e n é o número total de nós. O modelo computacional usa uma matriz de adjacências A , contendo os elementos $a_{i,j}$, em que $(a_{i,j} \in \{0, 1\})$. A matriz A é exibida na Eq. (5.3). Se $a_{i,j} = 0$, os nós da rede i e j não estão conectados, caso contrário eles estarão conectados, $(a_{i,j} = 1)$, e o enlace correspondente usa amplificadores ópticos predefinidos para compensar as perdas do enlace.

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{bmatrix}. \quad (5.3)$$

Nesta tese também foram considerados enlaces bidirecionais entre pares de nós, *i.e.* $a_{i,j} = a_{j,i}$. Visto que a matriz A é simétrica, $(a_{i,j} = a_{j,i})$, é possível representar esta matriz de forma “plana” pela justaposição, lado a lado, dos valores que estão acima da diagonal principal de A . Portanto, um vetor contendo $K = \frac{n^2-n}{2}$ entradas é usado. Este vetor é definido como $\vec{X} = \{x_i\}$, em que $(i \in 1, 2, \dots, K+2)$. O elemento x_{K+1} representa o tipo de ROADM usado e o elemento x_{K+2} representa o número máximo de comprimentos de onda por enlace. $x_{K+1} \in \{\text{ROADM}_0, \dots, \text{ROADM}_5\}$; em que ROADM_0 e ROADM_5 correspondem ao primeiro e ao último modelo de comutadores ópticos disponíveis. $x_{K+2} \in \{W_{min}, \dots, W_{max}\}$; em que W_{min} e W_{max} são o número mínimo e máximo de comprimentos de onda por enlace. Além disso, é incluído o requisito de que a topologia física da rede represente um grafo conectado. Esta restrição é necessária para garantir que seja possível que qualquer nó arbitrário na rede i possa se comunicar com qualquer outro nó j . Finalmente, uma descrição formal do problema é fornecida pelo enunciado a seguir.

Descrição do problema: Dado um conjunto de variáveis de decisão representado por

$$\vec{X} = [a_{1,2} \ a_{1,3} \ \dots \ a_{n-1,n} \ \delta \ w]. \quad (5.4)$$

$$\text{Minimize} \begin{cases} f_1(\vec{X}) = \text{CAPEX} \\ f_2(\vec{X}) = \text{PB} \end{cases}, \quad (5.5)$$

com a restrição de que

$$g(\vec{X}) = \lambda_{N-1} > 0, \quad (5.6)$$

em que λ_{N-1} é a conectividade algébrica do grafo e $\lambda_{N-1} > 0$ garante que a topologia física da rede óptica é um grafo conectado.

5.3 Um Novo Método para Inicialização de Projetos de Redes Ópticas

5.3.1 Descrição

É proposto um novo método para proposição de soluções iniciais para projetos de redes ópticas, denominado WS-T (Watts-Strogatz guiado por Tráfego). Este método será usado como um operador de inicialização em um algoritmo evolucionário de otimização multiobjetivo na Seção 5.4.

O algoritmo proposto nesta Tese é baseado no procedimento de Watts-Strogatz (WS), mas o algoritmo original é alterado para considerar uma matriz de tráfego e evitar grafos desconectados. A hipótese usada para escolha do modelo WS como ponto de partida para uma heurística de geração de redes ópticas está relacionada com as propriedades que este modelo confere às redes geradas. Conforme visto no Capítulo 3, as redes geradas pelo modelo WS possuem um comprimento médio dos caminhos relativamente baixo mesmo para redes com poucos enlaces. Além disso, o modelo WS parte de uma rede regular, o que confere uma topologia com boa capacidade na distribuição de tráfego para redes com multiplexação por comprimento de onda. A função usada para propor um novo projeto de rede óptica é descrita no Algoritmo 5. No passo 2 é escolhida uma densidade d entre uma densidade mínima e uma densidade máxima. No passo 3 a densidade é usada para calcular o número de enlaces da rede. O passo 4 corresponde à construção de uma rede k -regular, em que o valor de k dependerá da densidade escolhida. Os passos 5 até 10 são usados para criar “atalhos” na rede, mudando a posição original dos enlaces da rede k -regular para novas posições que potencialmente possuem maior demanda por tráfego. Os passos 11 até 13 são usados para acrescentar novos enlaces, além da quantidade já fornecida pela rede k -regular. Os passos 11 até 13 também consideram a matriz de tráfego. A topologia física criada é finalmente retornada no passo 14.

O Algoritmo 6 foi projetado para adicionar um novo enlace ao grafo se um número aleatório, uniformemente distribuído entre 0 e 1, é maior do que uma posição em uma matriz de tráfego normalizada. Por exemplo, se for esperado um grande volume de tráfego entre os nós i e j , $T[i][j]$ possuirá um valor próximo de 1,0 e este tipo de enlace será selecionado mais frequentemente. Este mecanismo é um tipo de conexão preferencial por tráfego. Neste procedimento, é impedida a criação de um enlace para o próprio nó considerado e enlaces que já estão presentes na rede.

Algoritmo 5 Algoritmo para construção eq topologia física de uma rede óptica.

```

1: função WS-T(nodes,  $d_{min}$ ,  $d_{max}$ ,  $T$ )
2:    $d \leftarrow rand() \cdot (d_{max} - d_{min}) + d_{min}$ 
3:    $nl \leftarrow d \cdot nodes(nodes - 1)/2$ 
4:    $A \leftarrow buildLimitedKRegular(n, d)$ 
5:   para  $link \leftarrow 0$ ;  $link < links.length$ ;  $link++$  faça
6:     se  $d[link.s] > 1 \wedge d[link.d] > 1 \wedge rand() > T[link.s][link.d]$  então
7:        $A[link.s][link.d] \leftarrow 0$ 
8:        $addLinkConsTraffic(A, nodes, T)$  ▷ Algoritmo 6
9:     fim se
10:  fim para
11:  enquanto  $links.length < nl$  faça
12:     $addLinkConsTraffic(A, nodes, T)$  ▷ Algoritmo 6
13:  fim enquanto
14:  retorne  $A$ 
15: fim função

```

Algoritmo 6 Algoritmo para acrescentar um enlace considerando uma matriz de tráfego.

```

1: procedimento ADDLINKCONSTRAFFIC( $A$ , nodes,  $T$ )
2:    $s \leftarrow rand(nodes)$ 
3:    $d \leftarrow rand(nodes)$ 
4:   enquanto  $s = d \vee A[s][d] = 1 \vee (s = link.s \wedge dst = link.d) \vee rand() > T[s][d]$  faça
5:      $s \leftarrow rand(nodes)$ 
6:      $d \leftarrow rand(nodes)$ 
7:   fim enquanto
8:    $A[s][d] \leftarrow 1$ 
9: fim procedimento

```

5.3.2 Arranjo Experimental

De modo a avaliar a proposta descrita, foi utilizado um conjunto de experimentos criados por meio de simulações. O foco dos experimentos realizados foi em projetar redes ópticas com diferentes números de nós e diferentes coberturas geográficas. As análises contemplam quatro diferentes distribuições de nós derivadas de redes conhecidas, a saber: rede NSFNet original, de 14 nós; rede MEDIANET [97], de 18 nós; rede ARNES [96], de 34 nós; e rede FLTG [95], de 41 nós. Estas redes foram descritas no Capítulo 4, no qual podem ser encontrados maiores detalhes sobre estas redes, incluindo os mapas da topologia de fibra. É importante lembrar que o modelo computacional adotado possui um vetor de variáveis de decisão que é proporcional à quantidade de nós da rede. Portanto, cada rede analisada exige um vetor de variáveis de decisão progressivamente maior, conforme ilustrado na Tabela 5.1.

Foi desenvolvida uma plataforma para execução de algoritmos evolucionários na linguagem de programação Java que implementa os algoritmos apresentados neste capítulo. As mesmas ferramentas mencionadas nos Capítulos 3 e 4 foram usadas nos experimentos deste capítulo

Tabela 5.1: Variáveis de decisão para planejamento das redes ópticas consideradas.

Rede	n	Topologia	δ e W	Total
NSFNet	14	91	2	93
MEDIANET	18	153	2	155
ARNES	34	561	2	563
FLTG	41	820	2	822

para permitir a execução de algoritmos relacionados com a Ciência das Redes e o treinamento de uma RNA.

Para simulação de redes ópticas também foi usado o simulador SIMTON. O simulador foi configurado para finalizar a estimativa de PB após a ocorrência de uma sequência de 500 bloqueios ou um total de 10^6 requisições de chamadas. A Qualidade de Transmissão (QoT) de cada caminho óptico é avaliada considerando os mesmos efeitos da camada física discutidos no Capítulo 4. Os parâmetros da rede usados nas simulações estão listados na Tabela 5.2.

Tabela 5.2: Parâmetros de simulação para o projeto das redes considerados no arranjo experimental.

Parâmetro	Valor	Definição
α	0,2 dB/km	Coefficiente de atenuação da fibra
Δt_{QoT}	0,10	Alargamento máximo permitido para o pulso óptico
Δf	100 GHz	Espaçamento entre canais
λ_i	1528,77 nm	Primeiro comprimento de onda da grade
λ_0	1450 nm	Zero de dispersão na fibra de transmissão
λ_{0RD}	1528,77 nm	Zero de dispersão residual
B_o	100 GHz	Largura de banda do filtro
B	10 Gb/s	Taxa de transmissão
D_{DCF}	-110 ps/km.nm	Coefficiente de dispersão da fibra de compensação
D_{PMD}	0,04 ps/ \sqrt{km}	Coefficiente de PMD
D_{TX}	4,5 ps/km.nm	Coefficiente de dispersão da fibra de transmissão
$Load$	100-300 erlangs	Carga de tráfego da rede
L_{Mux}	3 dB	Perda no multiplexador
L_{Demux}	3 dB	Perda no demultiplexador
L_{Switch}	3 dB	Perda no comutador óptico
NF	5 dB	Figura de ruído do amplificador
P_{Sat}	20 dB	Potência de saturação do amplificador
P_{TX}	3 dBm	Potência de transmissão
S_{DCF}	-1,87 ps/km.nm ²	<i>Slope</i> da fibra de compensação
S_{TX}	0,045 ps/km.nm ²	<i>Slope</i> da fibra de transmissão
$OSNR_{in}$	30 dB	Relação sinal-ruído óptica na entrada
$OSNR_{QoS}$	17 dB	OSNR mínima no receptor
W_{MIN}	4	Número mínimo de comprimentos de onda por enlace
W_{MAX}	40	Número máximo de comprimentos de onda por enlace

A proposta apresentada nesta Seção foi avaliada em cenários de tráfego uniforme e não-uniforme. Para o cenário não-uniforme da rede NSFNet, foi usada uma versão bidirecional da matriz de tráfego contida em [127]. De modo a criar a matriz bidirecional foi calculada uma média do tráfego ascendente e descendente. Para as outras três redes consideradas não estão disponíveis os dados referentes ao tráfego real e nestes casos foi utilizado um modelo

gravitacional para geração de uma matriz de tráfego [128]. O modelo proposto em [128] calcula o tráfego $L_{i,j}$ entre dois nós i e j de acordo com

$$L_{i,j} = K \frac{P_i P_j}{d_{i,j}^2}. \quad (5.7)$$

em que P_i é a população associada ao nó i , P_j é a população associada ao nó j , $d_{i,j}^2$ é a distância física entre i e j e K é um fator de escala. Nesta Tese, $d_{i,j}^2$ foi calculada usando a distância Euclidiana entre as coordenadas geográficas dos nós. K foi escolhido de modo a escalar a matriz para a carga de tráfego dinâmica, em erlangs, pretendida em cada estudo. As quatro matrizes de tráfego estão disponíveis no Apêndice B.

O algoritmo WS-T proposto neste capítulo foi comparado com outros cinco procedimentos. Os dois primeiros procedimentos foram propostos em trabalhos anteriores sobre planejamento de redes ópticas, a saber: CHAVES [49] e ARAUJO [22]. CHAVES [49] consiste em selecionar, por meio de uma distribuição uniforme, cada uma das posições do vetor \vec{X} que representa a configuração da rede. Ou seja, em CHAVES [49] há uma probabilidade de $\frac{1}{2}$ em estabelecer um enlace de fibra óptica na rede, uma probabilidade de $\frac{1}{6}$ de escolher um dos 6 modelos de comutadores ópticos e uma probabilidade de $\frac{1}{40}$ para seleção do número máximo de comprimentos de onda, por enlace. Considerando que a probabilidade adotada para estabelecimento de conexões em CHAVES [49] promove a geração de redes excessivamente conectadas, o procedimento usado em ARAUJO [22] consiste na redução da probabilidade de conexão de enlaces de fibra óptica para 0,30. Os outros três procedimentos comparados com WS-T são adaptações de algoritmos para geração de grafos conhecidos na literatura de Ciência das Redes (ER, BA e WS). Contudo, a probabilidade de estabelecimento de enlaces para estes modelos não é fixa. É definido um valor mínimo e máximo de densidade de enlaces para a rede e um valor de densidade é aleatoriamente definido para cada configuração de rede. A densidade é usada como um direcionador para o cálculo dos parâmetros específicos de cada modelo, tais como o Δm para redes BA, k para as redes WS e p para as redes ER. Para todos os inicializadores considerados, o modelo de ROADM e o número máximo de comprimentos de onda são selecionados por meio de uma distribuição uniforme.

Um experimento para análise dos algoritmos de inicialização consiste na criação de 100 soluções para cada operador de inicialização e em seguida a remoção de soluções dominadas. Foram realizadas 20 execuções independentes do experimento descrito para cada algoritmo e eles foram comparados usando o indicador de qualidade *coverage set*. O objetivo desta análise é verificar qual procedimento fornece as melhores sementes para serem usadas como propostas iniciais para algoritmos de planejamento de redes ópticas.

Todos os experimentos foram realizados em um computador com um processador Intel i5 e com 4 GB de memória RAM.

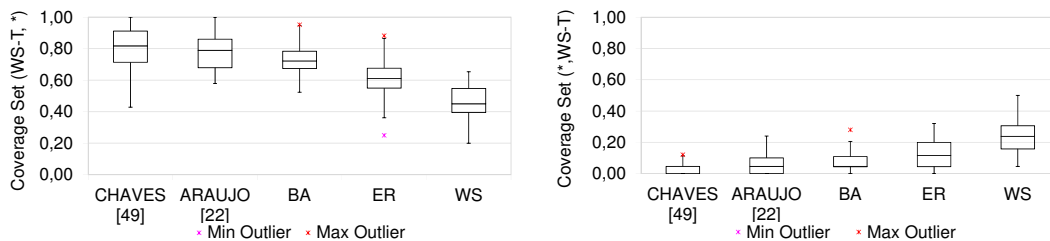
5.3.3 Resultados

Esta subseção está dividida em duas partes. A primeira parte contém uma discussão detalhada sobre os resultados da rede NSFNet, incluindo gráficos de *box-plot* das métricas de qualidade dos algoritmos e figuras com as curvas de frente de Pareto. A segunda parte sumariza os resultados para as outras três redes de forma tabulada e ilustra os resultados individuais de cada cenário por meio de frentes de Pareto.

Resultados para NSFNet

As Figuras 5.5 e 5.6 mostram os resultados do indicador *coverage set* para o procedimento WS-T e para outros métodos de inicialização para o caso NSFNet. Os gráficos de $CS(WS-T, *)$ fornecem a informação sobre o percentual de soluções de cada método que são dominadas pelo método WS-T. Por outro lado, os gráficos de $CS(*, WS-T)$ fornecem a informação sobre o percentual de soluções dos outros métodos que dominam as soluções do método WS-T. Portanto, é desejável que $CS(WS-T, *)$ seja acima de 0,5 e que $CS(*, WS-T)$ seja abaixo de 0,5.

De acordo com a Figura 5.5(a), as soluções de WS-T dominam mais de 60% das soluções de todos os outros algoritmos, exceto para o WS original, no qual $CS \cong 0,5$ (considerando uma comparação da mediana). Por outro lado, a Figura 5.5(b) mostra que as soluções da proposta WS-T são dominadas por menos do que 23% de todos os outros métodos, incluindo o método WS original.



(a) $CS(WS - T, *)$.

(b) $CS(*, WS - T)$.

Figura 5.5: Comparação de CS para diferentes algoritmos de inicialização, para o cenário de tráfego uniforme e para o caso NSFNet.

De acordo com a Figura 5.6(a), as soluções fornecidas por WS-T dominam mais de 76% das soluções de todos os outros métodos e dominam 100% das soluções obtidas por CHAVES [49] e ARAUJO [22], que são os melhores resultados na literatura sobre projeto de redes ópticas considerando uma modelagem do problema semelhante à adotada nesta Tese. Por outro lado, a Figura 5.6(b) mostra que soluções do algoritmo WS-T são dominadas por menos de 5% das soluções do algoritmo WS original e $CS(*, WS-T) = 0$ para todos os outros casos (excetuando-se os *outliers*).

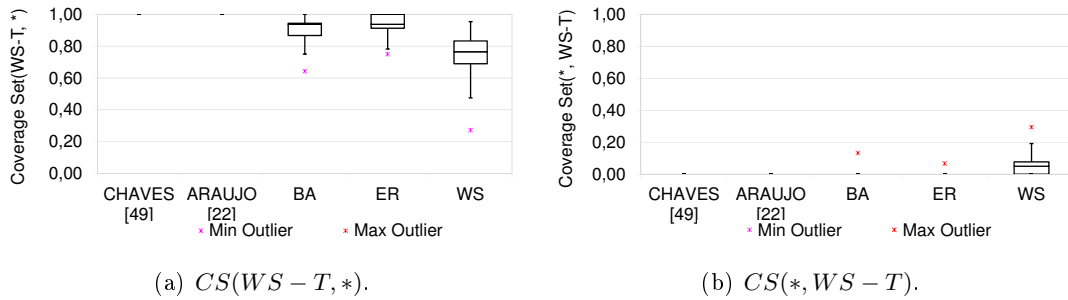


Figura 5.6: Comparação de CS para diferentes algoritmos de inicialização, para o cenário de tráfego não-uniforme e para o caso NSFNet.

Finalmente, a Figura 5.7 apresenta frentes de Pareto típicas para o procedimento WS-T e para duas propostas encontradas na literatura sobre projeto de redes ópticas. As frentes de Pareto ajudam na identificação da vantagem real no uso dos algoritmos, principalmente considerando a perspectiva do decisor humano, que escolherá um dos pontos da frente de Pareto para ser implantada. A contribuição da proposta WS-T como operador de inicialização em um sistema de planejamento de redes é clara para ambos os cenários de tráfego uniforme e não-uniforme, e neste último a vantagem é inquestionável sobre os outros métodos. Uma análise detalhada de alguns pontos da Figura 5.7(b) revela o benefício prático do novo método proposto em termos de redução de custo ou de melhoria de desempenho das redes. Por exemplo, quando são consideradas as soluções de menor PB de ARAUJO [22] e de WS-T na Figura 5.7(b), é possível identificar que o método proposto obtém uma rede que apresenta $PB = 8,6 \cdot 10^{-4}$ com o custo equivalente a uma rede de $PB = 3,7 \cdot 10^{-2}$, esta última obtida por ARAUJO [22]. Além disso, se for fixado o nível de desempenho da rede em $PB = 3,7 \cdot 10^{-2}$, a Figura 5.7(b) sugere que o método WS-T oferece uma rede com custo 30% menor.

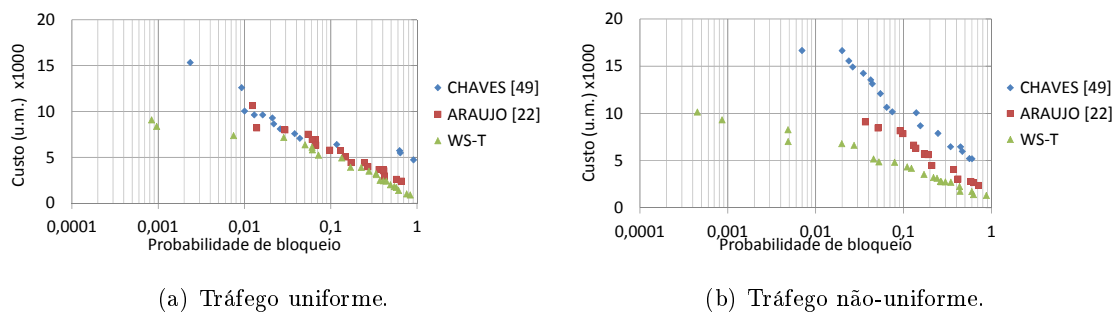


Figura 5.7: Frentes de Pareto obtidas para diferentes algoritmos de inicialização de projetos de redes ópticas para o caso NSFNet.

Resultados para ARNES, FLTG e MEDIANET

A Tabela 5.3 apresenta o valor médio do indicador *coverage set* para cada operador de inicialização, para o cenário de tráfego uniforme. A Tabela 5.4 apresenta a mesma informação para o cenário de tráfego não-uniforme.

Tabela 5.3: Média do indicador *coverage set*, para tráfego uniforme e para os cenários FLTG, ARNES e MEDIANET.

Cenário	$CS(WS-T, *)$					$CS(*, WS-T)$				
	[49]	[22]	BA	ER	WS	[49]	[22]	BA	ER	WS
FLTG	1,00	1,00	0,71	0,54	0,33	0,00	0,00	0,11	0,18	0,34
ARNES	1,00	0,94	0,27	0,26	0,41	0,00	0,01	0,46	0,38	0,30
MEDIANET	0,92	0,91	0,04	0,32	0,34	0,01	0,03	0,54	0,26	0,30

Tabela 5.4: Média do indicador *coverage set*, para tráfego não-uniforme e para os cenários FLTG, ARNES e MEDIANET.

Cenário	$CS(WS-T, *)$					$CS(*, WS-T)$				
	[49]	[22]	BA	ER	WS	[49]	[22]	BA	ER	WS
FLTG	1,00	0,97	0,85	0,88	0,77	0,00	0,00	0,08	0,05	0,20
ARNES	1,00	1,00	1,00	0,99	0,92	0,00	0,00	0,00	0,00	0,02
MEDIANET	1,00	1,00	0,81	0,96	0,87	0,00	0,00	0,12	0,02	0,06

Na Tabela 5.3 e na Tabela 5.4, os valores em negrito representam as situações nas quais o procedimento WS-T é mais vantajoso quando comparado aos demais. Uma visão geral dos valores de *CS* apresentado nestas tabelas revelam que WS-T é mais vantajoso para todas as comparações envolvendo tráfego não-uniforme (Tabela 5.4). Para o caso uniforme, o método proposto é superior aos dois métodos já publicados na literatura, ou seja, WS-T é sempre mais vantajoso quando comparado com CHAVES [49] e ARAUJO [22]. Para redes FLTG e para o caso uniforme, o método WS-T é superior aos métodos BA e ER. Para o caso MEDIANET e para tráfego uniforme, o procedimento BA se mostra mais vantajoso. Finalmente, de acordo com o indicador *CS*, para o caso ARNES e para tráfego uniforme, é indiferente a escolha dos procedimentos BA, ER, WS e WS-T. De forma quantitativa, se o critério de decisão para uso dos métodos analisados é $CS(WS-T, *) > 0,5$ e $CS(*, WS-T) < 0,5$, percebe-se que o método WS-T é melhor em 23 casos, é pior em 1 caso e é indiferente em 5 dos casos considerados.

Alguns dos resultados de *CS* merecem destaque. A superioridade do método WS-T sobre CHAVES [49] e ARAUJO [22] deve-se principalmente ao fato de estes dois últimos usarem uma probabilidade de conexão de enlaces fixa, enquanto que o procedimento WS-T escolhe uma probabilidade de religação em uma faixa mais adequada de grau médio para redes de transporte. Por este motivo, os demais inicializadores propostos nesta Tese, baseados em BA, ER e WS, também são mais promissores quando comparado com CHAVES [49] e ARAUJO [22]. O método WS-T se mostrou muito superior para os casos de tráfego não-uniforme, perante todos os demais métodos, principalmente devido ao procedimento de religação orientado pela

matriz de tráfego. O procedimento de BA se mostrou uma alternativa interessante no caso uniforme da MEDIANET pois este método promove a criação de concentradores (*hubs*) e uma topologia baseada em concentradores geralmente apresenta CAPEX mais baixo. Perceba que, uma modelagem do problema que incorpora a robustez da rede como um terceiro objetivo, além de CAPEX e PB, penalizaria os valores de CS para o método BA, pois redes BA são vulneráveis a falhas em nós concentradores [64].

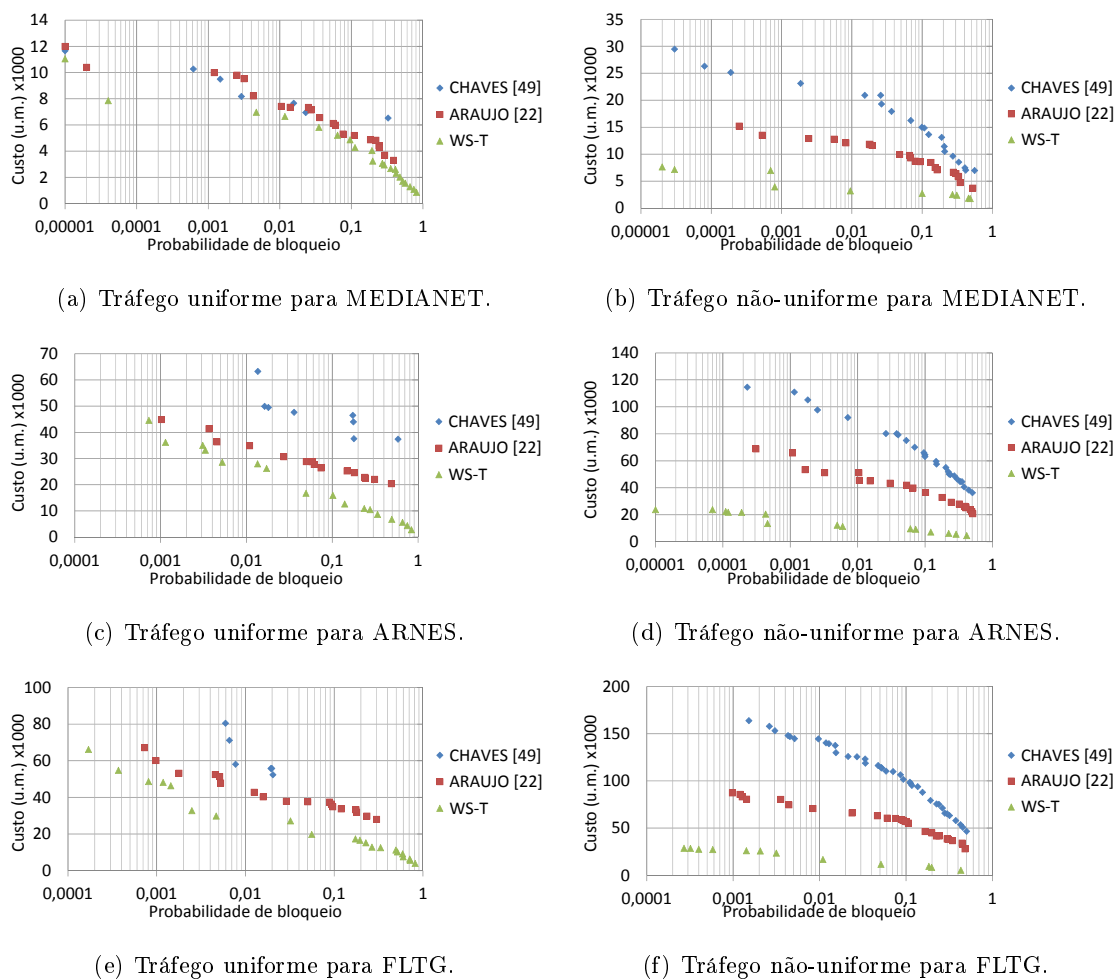


Figura 5.8: Frentes de Pareto obtidas para diferentes algoritmos de inicialização de projetos de redes ópticas para MEDIANET, ARNES e FLTG.

A Figura 5.8 apresenta as frentes de Pareto para redes MEDIANET, ARNES e FLTG e para os cenários de tráfego uniforme e não-uniforme. Uma visão geral da Figura 5.8 permite concluir que o método WS-T de fato é superior em todos os casos considerados, mas é especialmente vantajoso para tráfego não-uniforme.

De acordo com a Figura 5.8(b), para redes MEDIANET e para o cenário não-uniforme, é possível obter redes que apresentam $10^{-4} < PB < 10^{-3}$ usando o método WS-T com

uma redução de custo de 70% em comparação com o método ARAUJO [22] e de 83% em comparação com o método CHAVES [49].

De acordo com a Figura 5.8(d), para redes ARNES e para o cenário não-uniforme, é possível obter redes que apresentam $10^{-4} < PB < 10^{-3}$ usando o método WS-T com uma redução de custo de 69% em comparação com o método ARAUJO [22] e de 81% em comparação com o método CHAVES [49]. Além disso, o método WS-T foi o único que obteve redes com $PB < 10^{-4}$ no exemplo apresentado na Figura 5.8(d).

Para redes FLTG e para o cenário não-uniforme, é possível obter redes que apresentam $10^{-3} < PB < 10^{-2}$ usando o método WS-T com uma redução de custo de 68% em comparação com o método ARAUJO [22] e de 84% em comparação com o método CHAVES [49].

Considerações Finais sobre o Operador de Inicialização

Nesta seção, o método WS-T foi aplicado para propor topologias físicas para quatro diferentes distribuições de nós e para os cenários de tráfego uniforme e não-uniforme.

Para o cenário de tráfego uniforme, e considerando o comportamento mais frequente para os quatro casos analisados, o método WS-T se mostrou mais promissor quando comparado com procedimentos disponíveis na literatura e se mostrou equivalente aos demais procedimentos propostos. Para o cenário de tráfego não-uniforme, o procedimento WS-T apresenta vantagem inquestionável sobre todos os demais procedimentos quando é considerado o compromisso entre custo e desempenho da rede. De acordo com os resultados obtidos, o método WS-T oferece uma redução de custo de até 70% quando comparado com a melhor proposta já publicada na literatura.

É importante mencionar que o método proposto nesta Tese é uma contribuição relevante principalmente se for considerado que o cenário de tráfego não-uniforme é o mais frequente em projetos de redes de transporte, pois na maioria dos casos são considerados aspectos socio-econômicos das localidades que representam os nós da rede para composição de demandas de tráfego. Apesar de o método proposto ser mais promissor quando comparado com os demais, é importante destacar que o uso exclusivo deste procedimento ainda não oferece configurações de rede otimizadas. Um processo de otimização deve complementar os algoritmos de geração de topologias físicas para obtenção de redes com menor custo e menor PB.

5.4 Cascata de Modelos Substitutos para Avaliação de Redes em Algoritmos de Otimização

5.4.1 Descrição

Embora os algoritmos propostos na Seção 5.3 sejam um primeiro passo na direção de uma ferramenta de planejamento, o uso exclusivo destes procedimentos ainda não permite a obten-

ção de redes otimizadas, ou seja, as soluções propostas na Seção 5.3 ainda não correspondem ao melhor compromisso em termos de CAPEX e PB.

Nesta seção é proposta uma nova abordagem evolucionária multiobjetiva para tratar o problema de projeto de redes ópticas. A principal contribuição desta proposta é introduzir o conceito de Cascata de Modelos Substitutos (CMS), que consiste em uma metodologia para obtenção de avaliações rápidas da PB de redes WDM durante o processo evolucionário. Além disso, o procedimento WS-T proposto na Seção 5.3 é usado como um operador de inicialização no algoritmo evolucionário.

A otimização assistida por modelos substitutos para projeto de redes ópticas ainda é uma área de investigação em aberto. A proposição de um modelo substituto baseado em RNA para estimativa de PB é um primeiro passo na direção do uso de uma abordagem desta natureza em processos de otimização. Contudo, o algoritmo de planejamento de rede deve prever a possibilidade de o estimador alternativo fornecer valores incorretos de PB para evitar que o processo de otimização seja comprometido. Por exemplo, considere uma configuração de rede que possui uma PB estimada por simulação na ordem de 10^{-4} e que diversos bloqueios nesta rede são devidos a algum efeito não capturado pela RNA, como FWM. Caso a PB desta configuração de rede seja avaliada pela RNA na ordem de 10^{-5} , esta configuração provavelmente dominaria diversas outras em termos de custo e o processo inteiro poderia convergir para um ponto incorreto do espaço de busca. Nesta Tese é proposto um mecanismo de controle dos modelos substitutos que leva em consideração a correlação entre a PB estimada e a faixa de custo esperada em um dado momento do processo de busca. De acordo com a classificação fornecida na Seção 5.1.4, a proposta desta Tese usa múltiplos modelos substitutos em uma abordagem baseada em indivíduos na qual a precisão das estimativas dos indivíduos é considerada para o gerenciamento do uso da função de aptidão original. Também foram realizados estudos sobre abordagens baseadas em gerações, mas o AE frequentemente é conduzido para falsos pontos ótimos quando a abordagem baseada em geração é usada.

Nesta Tese, a função de custo usada é a mesma adotada em [21], mas o mecanismo para estimativa de PB é completamente modificado. No novo método de estimativa de PB proposto nesta Tese, o tempo necessário para estimar a PB de uma rede é bem menor e a precisão fornecida é semelhante à de trabalhos anteriores. Uma estimativa de PB geralmente é obtida por meio de simulações tipo Monte Carlo, em que um grande número de requisições de chamadas (RC) é considerado. De modo a obter uma estimativa de PB com precisão para redes com $PB > 10^{-5}$ são usadas $RC > 10^7$. Contudo, esta abordagem requer de alguns segundos até minutos em um computador típico. O método CMS proposto nesta Tese é formado por um conjunto de modelos substitutos em que cada elemento do conjunto é capaz de estimar a PB de redes apresentando um compromisso entre precisão da estimativa e tempo de resposta. O método CSM escolhe dinamicamente o modelo substituto que possui a precisão mínima requerida para a rede considerada, evitando o tempo adicional de simulação que seria

necessário no método convencional.

Tabela 5.5: Modelos Substitutos para estimativa de PB de redes ópticas.

Modelo	Descrição	Intervalo preferencial de PB
1	Modelo baseado em RNA	$0,10 < PB < 1,00$
2	Simulações de rede com $RC = 2.000$	$0,01 < PB < 0,10$
3	Simulações de rede com $RC = 10.000$	$10^{-3} < PB < 0,01$
4	Simulações de rede com $RC = 100.000$	$10^{-4} < PB < 10^{-3}$
5	Simulações de rede com $RC = 1.000.000$	$10^{-5} < PB < 10^{-4}$

A Tabela 5.5 apresenta um exemplo de conjunto de modelos substitutos usados na proposta desta seção para estimativa de PB. O primeiro modelo substituto é uma RNA que usa como variáveis de entrada métricas topológicas e características da camada física. Ou seja, a RNA equivale à proposta de estimativa de PB apresentada no Capítulo 4. Modelos substitutos geralmente são construídos usando a representação completa de uma solução do problema como entrada para o modelo, contudo o problema considerado nesta Tese possui um grande número de entradas relacionadas com a topologia da rede. Portanto, um aspecto chave do método proposto é que a camada de entrada da RNA é um resumo da topologia física, permitindo a extensão do trabalho para redes de tamanhos arbitrários. Logo, a solução possui escalabilidade com relação ao número de nós da rede. Na Tabela 5.5, os modelos 2-5 são baseados em simulações de requisições de chamadas à rede, mas fornecem uma alternativa ao uso de uma quantidade muito grande de chamadas ($RC \geq 10^7$), evitando simulações de chamadas desnecessárias. O conceito de intervalo preferencial de PB significa que o modelo substituto garante uma estimativa precisa para redes que possuem PB neste intervalo. Contudo, qualquer modelo substituto pode estimar uma WRON que apresenta PB fora do intervalo preferencial, pois um procedimento dinâmico é usado para estimativa da probabilidade de erro no cálculo da PB.

Com o objetivo de aceitar ou rejeitar a estimativa fornecida por um modelo substituto “mais barato” é usado um modelo simples que detecta uma precisão inadequada nas avaliações usando previsões por faixa de custo das redes. Este procedimento está descrito no Algoritmo 7 e usa como procedimento auxiliar o Algoritmo 8. Uma explicação detalhada destes procedimentos é fornecida a seguir. O espaço de objetivos é dividido em NC bandas de custo. São utilizadas médias móveis e desvios padrões da probabilidade de bloqueio em cada faixa para decidir se as avaliações devem ser aceitas ou não. Na Figura 5.9, a faixa de custo i possui uma média móvel histórica de PB igual a \overline{PB}_i e um desvio padrão igual a σ_i . Portanto, a regra de decisão para avaliações de novas soluções é: aceite avaliações se $\overline{PB}_i - n\sigma_i < PB < \overline{PB}_i + n\sigma_i$, e rejeite as avaliações caso contrário. Cada vez que uma estimativa é rejeitada, o fluxo é redirecionado para uma nova avaliação por um modelo mais custoso e mais preciso, até que a avaliação seja aceita. A verificação da posição de uma solução de acordo com as faixas

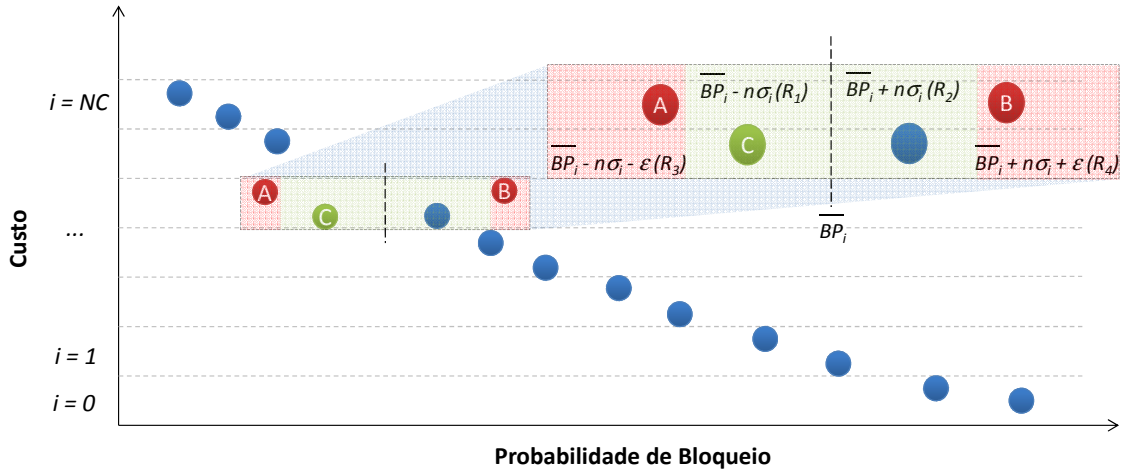


Figura 5.9: Ilustração da metodologia proposta para tratamento da cascata de modelos substitutos.

de custo é feita pelo Algoritmo 8 e a posição encontrada é usada nos testes das linhas 7 e 10 do Algoritmo 7. É usada uma regra de atualização dinâmica para \overline{PB}_i e σ_i (linha 16 do Algoritmo 7). \overline{PB}_i é atualizado de acordo com

$$\overline{PB}_i(t) = \alpha \overline{PB}_i(t-1) + (1 - \alpha) PB, \quad (5.8)$$

em que $\overline{PB}_i(t-1)$ é a média de PB para a faixa de custo i antes da ocorrência desta avaliação; PB é uma estimativa de PB para uma nova WRON que possui um custo localizado na faixa i ; e α é uma constante usada para ajustar a importância relativa de novas avaliações sobre as avaliações antigas de PB. É desejável que $\alpha < 0,5$ para fornecer um peso maior para avaliações recentes e garantir que o processo evolucionário não seja retardado. σ_i é atualizado de acordo com

$$\sigma_i(t) = \alpha \sigma_i(t-1) + (1 - \alpha) |\overline{PB}_i(t) - PB|, \quad (5.9)$$

em que $\sigma_i(t-1)$ é o desvio padrão de PB para a faixa de custo i anterior a esta avaliação de rede.

O procedimento completo para o projeto de redes WDM é apresentado no Algoritmo 9. Foi usada a estrutura básica do NSGAI, mas o algoritmo usa os procedimentos descritos previamente (Algoritmos 5 até 8). O procedimento *criarModelo Substituto()* cria um conjunto de redes WDM geradas aleatoriamente e treina uma RNA para avaliação de redes WDM conforme explicado no Capítulo 4. O procedimento de seleção de pais (linha 11) é o método tradicional de torneio binário. A linha 12 está relacionada com o cruzamento uniforme [129].

Algoritmo 7 Algoritmo para tratar uma Cascata de Modelos Substitutos para estimativa de PB de redes ópticas.

```

1: procedimento AVALIARSOLUCAOPORCMS(rede)
2:    $A \leftarrow rede.A$ 
3:    $W \leftarrow rede.W$ 
4:    $\delta \leftarrow rede.ROADM$ 
5:    $PB \leftarrow calcularSaidaRna(A, W, \delta)$ 
6:    $custo \leftarrow calcularCustoRede(A, W, \delta)$ 
7:   se  $PB < left(1) \vee posEmBandaCusto(PB, custo) = IN$  então
8:     para  $id \leftarrow 2; id \leq 5; id ++$  faça
9:        $pos \leftarrow posEmBandaCusto(PB, custo)$ 
10:      se  $(PB < left(id)) \wedge (pos = NULL \vee pos = LEFT)$  então
11:         $PB \leftarrow simularPB(A, W, \delta, NC(id))$ 
12:      fim se
13:    fim para
14:  fim se
15:  se  $PB < left(i) \wedge posEmBandaCusto(PB, custo) \neq RIGHT$  então
16:     $atualizarFaixaCusto(PB, custo)$  ▷ De acordo com Eqs. (5.7) e (5.8)
17:  fim se
18:   $rede.pb \leftarrow PB$ 
19:   $rede.custo \leftarrow custo$ 
20: fim procedimento

```

Algoritmo 8 Algoritmo para avaliar a posição de uma solução de acordo com a sua faixa de custo.

```

1: função POSEMBANDACUSTO( $PB, custo, n$ )
2:    $PB_i \leftarrow pbMedioPorFaixaCusto(custo)$ 
3:    $\sigma_i \leftarrow desvioMedioPorFaixaCusto(custo)$ 
4:   se  $PB > PB_i + n\sigma_i$  então
5:      $value \leftarrow RIGHT$  ▷ Está fora da faixa pela direita
6:   senão se  $PB < PB_i - n\sigma_i$  então
7:      $value \leftarrow LEFT$  ▷ Está fora da faixa pela esquerda
8:   senão se  $PB < PB_i + n\sigma_i$  AND  $PB > PB_i - n\sigma_i$  então
9:      $value \leftarrow IN$  ▷ Está dentro da faixa
10:  senão
11:     $value \leftarrow NULL$  ▷ É a primeira avaliação nesta faixa
12:  fim se
13:  return  $value$ 
14: fim função

```

A linha 13 usa um procedimento de mutação simples que troca 0's e 1's de acordo com P_M para as posições da matriz de adjacências. A mutação também altera W e $ROADM$ por um valor gerado aleatoriamente também respeitando o valor de P_M . A linha 22 usa uma versão modificada do operador “Distância da multidão” que considera $\log_{10} \Delta PB$ de modo a criar uma pressão adicional para manter redes WDM na região $10^{-5} < PB < 10^{-2}$.

Algoritmo 9 O novo algoritmo evolucionário proposto para planejamento de redes ópticas.

```

1: procedimento PTDOPTICALNETWORKS( $n, d_{min}, d_{max}, T, P, P_M, P_C, W_{min}, W_{max},$ 
    $ROADM_{max}, I_{max}$ )
2:   criarModeloSubstituto()           ▷ treinamento de RNA para um conjunto de redes
3:   para  $s \leftarrow 0; s < P; s ++$  faça
4:      $pop[s].A \leftarrow WS-T(n, d_{min}, d_{max}, T)$            ▷ Algoritmo 5
5:      $pop[s].W \leftarrow rand(W_{max} - W_{min}) + W_{min}$ 
6:      $pop[s].ROADM \leftarrow rand(ROADM_{max})$ 
7:     avaliarSolucaoPorCSM(population[s])           ▷ Algoritmo 7
8:   fim para
9:   para  $i \leftarrow 0; i < I_{max}; i ++$  faça           ▷ Processo evolucionário
10:    para  $s \leftarrow 0; s < P; s \leftarrow s + 2$  faça
11:     Obtenha os descendentes offspring1 e offspring2
12:     Faça cruzamento de offspring1 e offspring2
13:     Faça mutação de offspring1 e offspring2
14:     avaliarSolucaoPorCSM(offspring1, offspring2)           ▷ Algoritmo 7
15:     Acrescentar soluções offspring1 e offspring2 em offsprings
16:    fim para
17:     $union \leftarrow pop + offsprings$ 
18:     $fronts \leftarrow buildParetoFronts(union)$ 
19:     $pop \leftarrow emptySet(P)$ 
20:     $front \leftarrow 0$ 
21:    enquanto  $pop.length < P$  AND  $front < fronts.length$  faça
22:     ordenarPorCrowdingDistanceLog(fronts[front])
23:     para  $s \leftarrow 0; s < fronts[front].length; s ++$  faça
24:       se  $pop.length < P$  então
25:         Acrescentar a solução  $fronts[front][s]$  em  $pop$ 
26:       fim se
27:     fim para
28:      $front \leftarrow front + 1$ 
29:   fim enquanto
30: fim para
31: fim procedimento

```

5.4.2 Arranjo Experimental

Os experimentos realizados para análise do Algoritmo 9 faz uso da plataforma de simulação já descrita na Seção 5.3.2.

Para simulação de redes ópticas também foi usado o simulador SIMTON. Com o objetivo de avaliar a robustez do controle de modelos substitutos, diversas configurações do simulador de rede foram usadas, variando o número máximo de chamadas, de acordo com o intervalo preferencial de uso de cada configuração. Para as análises de tráfego uniforme, o SIMTON foi usado em conjunto com a RNA e o simulador foi configurado para finalizar a estimativa de PB após um total de RC requisições de chamadas, de acordo com a Tabela 5.5. Para o cenário de tráfego não-uniforme, foi avaliado o desempenho da mesma arquitetura de RNA usada para o caso uniforme, mas os erros de estimativa não foram aceitáveis e no caso não-uniforme foi usado apenas o SIMTON. O estudo de projeto de rede em cenário não-uniforme permite a avaliação da novo procedimento de inicialização em promover bons compromissos de projeto de redes, mesmo após várias iterações do processo evolucionário. Além disso, foram consideradas duas diferentes configurações de efeitos da camada física. Para um estudo inicial, realizado sobre a rede NSFNet foram ativados todos os efeitos da camada física disponíveis no SIMTON. Para análise das redes ARNES, FLTG e MEDIANET, o efeito FWM foi desconsiderado. A avaliação de FWM é computacionalmente custosa e não é capturada por modelos substitutos baseados em RNA, portanto, o uso destas duas configurações permite uma avaliação mais detalhada sobre o benefício da proposta. Vale a pena destacar que qualquer outro simulador de eventos discretos poderia ter sido usado para simulação de uma rede óptica sem comprometer os resultados obtidos neste capítulo.

Foi analisado o impacto dos parâmetros α e NC para a melhoria do funcionamento da cascata de modelos substitutos e os melhores valores encontrados para o projeto da rede NSFNet foram $\alpha = 0,6$ e $NC = 400$. I_{max} , P , P_M e P_C já foram estudados em trabalhos anteriores [22] e aqui serão adotados os mesmos valores. A densidade mínima aceita para a rede (d_{min}) está relacionada com uma rede em anel ($d_{min} = 1/(n - 1)$) e d_{max} foi escolhida por adequação a redes reais. A Tabela 5.6 resume os parâmetros adotados para avaliação do algoritmo evolucionário.

Tabela 5.6: Parâmetros do algoritmo evolucionário usado para projeto de redes.

Parâmetro	Valor	Definição
n	14, 18, 34 e 41	Número de nós para cada cenário (vide Tabela 5.1)
d_{min}	0,08	Densidade mínima
d_{max}	0,60	Densidade máxima
P	50	Tamanho da população
P_M	$\frac{1}{(n^2 - n)/2 + 2}$	Probabilidade de mutação
P_C	1,00	Probabilidade de cruzamento
I_{max}	1.000	Número máximo de iterações
α	0,6	Fator de amortecimento para atualização de $\overline{PB}_i(t)$ e $\sigma_i(t)$
NC	400	Número de bandas de custo

Com o objetivo de avaliar o algoritmo evolucionário proposto foram considerados dois tipos de experimentos. O primeiro tipo de experimento consiste na evolução de soluções por uma quantidade de tempo previamente delimitada e relacionada com 1.000 avaliações

de aptidão (*fitness*) do método tradicional. Esta quantidade de avaliações corresponde a aproximadamente 15 minutos no ambiente de execução considerado para uma rede de 14 nós e com todos os efeitos da camada física ativados. O segundo tipo de experimento consiste na execução do processo evolucionário até que $\Delta HV \leq 10^{-6}$. Ou seja, o primeiro experimento visa identificar a qualidade de soluções para um tempo de execução máximo pré-definido e o segundo tipo visa identificar o tempo necessário para a convergência do algoritmo. O algoritmo é comparado com a proposta do estudo apresentado em [99]. A qualidade das soluções é avaliada de acordo com as métricas *CS*, *HV* e *MS*.

5.4.3 Resultados Obtidos

Esta subseção será dividida em duas partes. Na primeira parte serão apresentados os resultados detalhados para a rede NSFNet, incluindo gráficos de *box-plot* das métricas de qualidade e figuras com as frentes de Pareto. Na segunda parte será apresentado um resumo dos resultados obtidos para o projeto das redes ARNES, FLTG e MEDIANET.

Na discussão dos resultados, EA-NS (*Evolutionary Algorithm-Network Simulator*) indica que a função objetivo $f_2(\vec{X})$ é avaliada apenas por um simulador de redes. O uso de simuladores de rede (EA-NS) para estimativa de PB é adotado em diversos estudos sobre planejamento de redes ópticas [49, 22, 99]. Por outro lado, EA-CSM (*Evolutionary Algorithm-Cascade of Surrogate Models*) representa o uso da nova proposta para avaliar a função objetivo $f_2(\vec{X})$, ou seja, o uso combinado de uma RNA e de um simulador de redes.

Resultados para NSFNet

A Figura 5.10 apresenta os indicadores de qualidade para o cenário uniforme, comparando um AE tradicional e a nova proposta baseada em CMS com o critério de parada por tempo fixo (equivalente a 1,000 avaliações da função de *fitness* original ou 15 minutos). A Figura 5.10(a) apresenta os resultados de *CS* para os dois algoritmos. De acordo com a Figura 5.10(a), após 15 minutos de execução de ambos os algoritmos, as soluções da nova abordagem dominam 98% das soluções do EA-NS e nenhuma solução do EA-NS domina as soluções obtidas pela nova proposta. A Figura 5.10(b) apresenta o HV para ambos os algoritmos. De acordo com a Figura 5.10(b), a nova proposta oferece um valor de HV, considerando a mediana dos resultados, que é 34% superior ao EA-NS. A Figura 5.10(c) apresenta o indicador MS para ambos os procedimentos. De acordo com a Figura 5.10(c), a nova proposta oferece um valor mediano de MS que é 37% superior ao fornecido pelo EA-NS. MS é especialmente útil para análise da capacidade dos algoritmos em fornecer redes com valores muito baixos de PB.

Também foi avaliado para o cenário uniforme o tempo necessário para a parada dos algoritmos quando nenhuma melhoria é percebida entre as gerações do algoritmo evolucionário. Se o critério de parada é $\Delta HV < 10^{-6}$, o EA-NS requer em média 1450 minutos para finalizar.

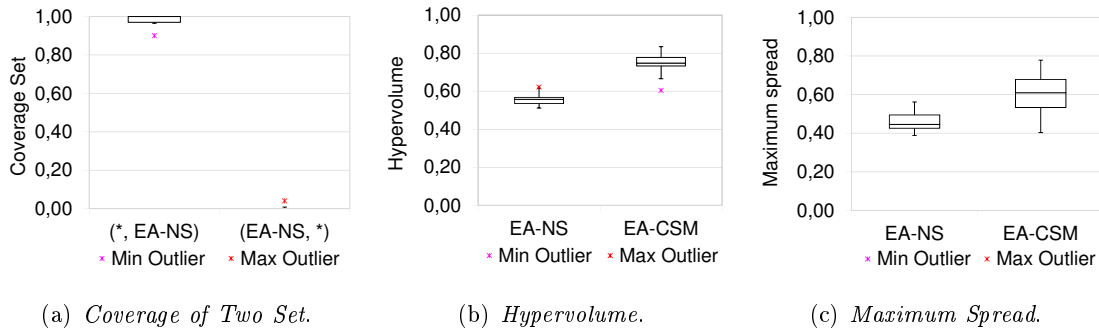


Figura 5.10: Indicadores de qualidade aplicados às frentes de Pareto obtidas após 15 minutos de execução para a nova proposta evolucionária e para o EA-NS. O cenário considerado foi o de tráfego uniforme para a NSFNet.

Por outro lado, a nova proposta baseada em Cascata de Modelos Substitutos requer apenas 226 minutos para finalizar. Nesta avaliação, ambos os algoritmos alcançam o mesmo nível de convergência e de espalhamento das soluções. Portanto, se uma completa convergência da frente de Pareto é necessária, a nova proposta reduz em 84% o tempo de execução ou 1224 minutos em valores absolutos.

A Figura 5.11 mostra, para o cenário de tráfego uniforme, três frentes de Pareto. P_{ref} corresponde a uma frente de Pareto de referência obtida pela combinação de diferentes execuções com critério de parada $\Delta HV < 10^{-6}$. EA-NS e EA-CSM são frentes de Pareto de um algoritmo tradicional e da nova proposta, respectivamente, para 15 minutos de execução. De acordo com a Figura 5.11, é possível concluir que a frente de Pareto fornecida pela nova proposta após um pequeno espaço de tempo é capaz de aproximar-se consideravelmente da frente de Pareto de referência. Este resultado sugere que é possível que um decisor use as sugestões de projetos de redes ópticas mesmo nos primeiros estágios do processo evolutivo, pois já serão soluções com um bom compromisso entre Custo e PB.

A Figura 5.12 apresenta os indicadores de qualidade para o cenário não-uniforme, comparando o EA-NS e a nova proposta baseada em CMS com o critério de parada por tempo fixo (equivalente a 1000 avaliações da função de *fitness* original ou 15 minutos). A Figura 5.12(a) apresenta os resultados de CS para os dois algoritmos. De acordo com a Figura 5.12(a), após 15 minutos de execução de ambos os algoritmos, as soluções do novo algoritmo proposto dominam 98% das soluções do EA-NS e nenhuma solução do EA-NS domina as soluções obtidas pela nova proposta. A Figura 5.12(b) apresenta o HV para ambos os algoritmos. De acordo com a Figura 5.12(b), a nova proposta oferece um valor mediano de HV que é 49% superior ao EA-NS. A Figura 5.12(c) apresenta o indicador MS para ambos os procedimentos. De acordo com a Figura 5.12(c), a nova proposta oferece um valor mediano de MS que é 39% superior ao fornecido pelo EA-NS.

Se o critério de parada para o cenário não-uniforme é $\Delta HV < 10^{-6}$, o EA-NS requer em

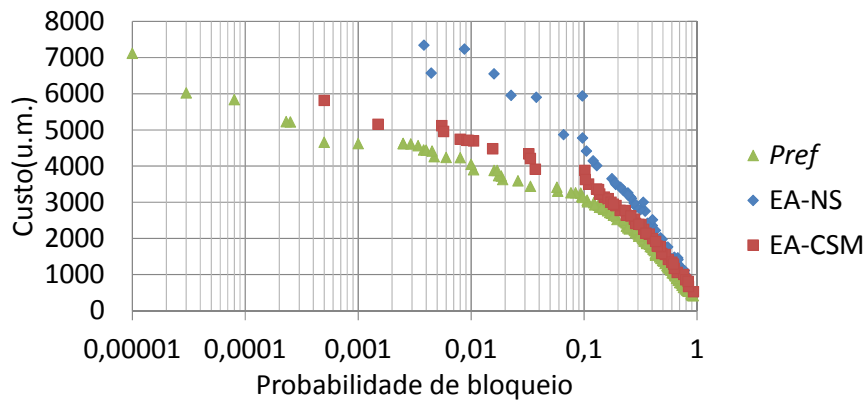


Figura 5.11: Frontes de Pareto obtidas após 15 minutos de execução do EA-NS e da nova proposta e uma frente de Pareto de referência. O cenário considerado é o de tráfego uniforme para a NSFNet.

média 968 minutos para finalizar. Por outro lado, a nova proposta baseada em Cascata de Modelos Substitutos requer apenas 112 minutos para finalizar. Portanto, se uma completa convergência da frente de Pareto é necessária, a nova proposta reduz em 88% o tempo de execução ou 856 minutos em valores absolutos para o cenário não-uniforme.

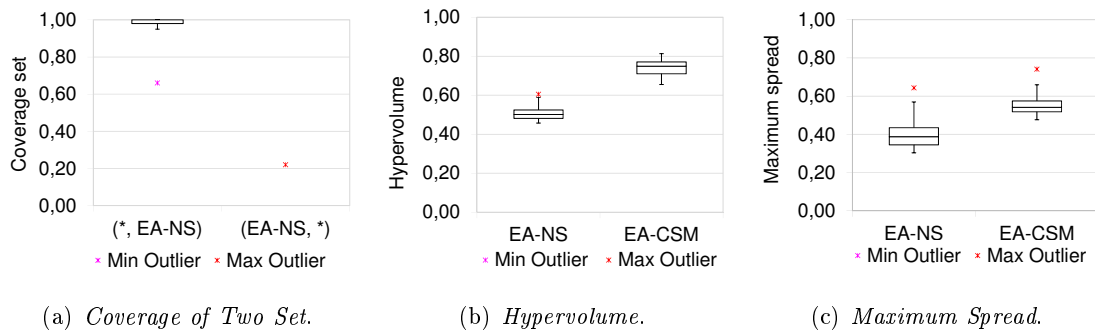


Figura 5.12: Indicadores de qualidade aplicados às frentes de Pareto obtidas após 15 minutos de execução para a nova proposta evolucionária e para o EA-NS. O cenário considerado foi o de tráfego não-uniforme para a NSFNet.

A Figura 5.13 mostra, para o cenário não-uniforme, três frentes de Pareto. P_{ref} corresponde a uma frente de Pareto de referência obtida pela combinação de diferentes execuções com critério de parada $\Delta HV < 10^{-6}$). EA-NS e EA-CSM são frentes de Pareto do EA tradicional e da nova proposta, respectivamente, para 15 minutos de execução. De acordo com a Figura 5.13, é possível concluir que a frente de Pareto fornecida pela nova proposta, após um pequeno intervalo de tempo, é capaz de aproximar-se consideravelmente da frente de Pareto de referência. Este resultado é ainda mais expressivo quando comparado ao que foi obtido

para o cenário de tráfego uniforme.

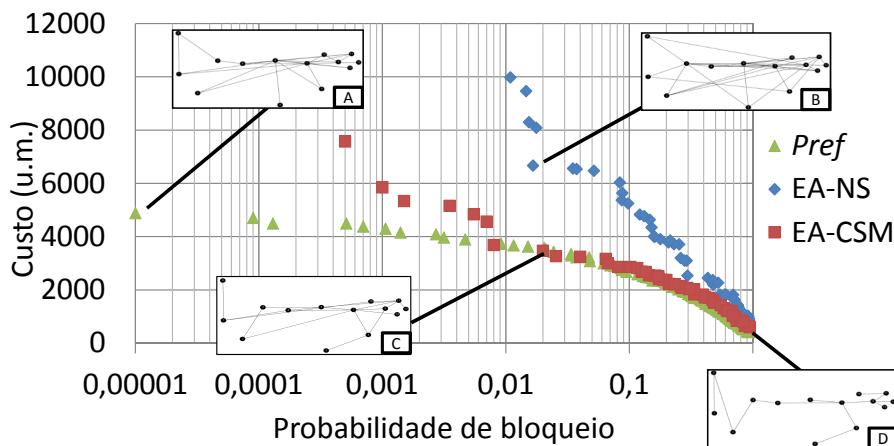


Figura 5.13: Frentes de Pareto obtidas após 15 minutos de execução do EA-NS e da nova proposta e uma frente de Pareto de referência. O cenário considerado é o de tráfego não-uniforme para a NSFNet.

A Figura 5.13 também destaca quatro pontos relacionados com diferentes configurações de redes WDM. A Tabela 5.7 contém informações detalhadas sobre estas redes. As configurações de rede *A* e *D*, contidas na frente de Pareto de referência, correspondem a configurações antagônicas. O projeto físico *A* corresponde a uma rede malhada e com PB muito baixa. Na configuração *A*, além da topologia de fibra ser suficientemente malhada, também foi usado um número de comprimentos de onda por enlace e dispositivo ROADM que favoreça o valor baixo de PB (vide Tabela 5.7). Em contrapartida, o projeto físico *D* está relacionado a uma rede de custo muito baixo, mas que é infactível para uso prático, pois possui $PB = 0,84$. Na configuração *D* há uma topologia de fibra mínima, são usados apenas 4 comprimentos de onda por enlace e a qualidade do ROADM é inferior quando comparada com as demais configurações. Portanto, as variáveis de decisão configuradas para *D* não fornecem uma PB aceitável para a matriz de tráfego uniforme dada e para uma carga de 200 erlangs. As configurações *B* e *C* estão na fronteira da região de interesse de uma rede óptica de transporte e oferecem uma comparação ilustrativa. O projeto físico *B*, obtido pelo EA-NS, oferece $PB = 1,66 \cdot 10^{-2}$ por meio de uma topologia de fibra muito malhada, uso de muitos comprimentos de onda e o melhor ROADM disponível. O projeto físico *C*, obtido pela nova proposta evolucionária, oferece $PB = 2,0 \cdot 10^{-2}$ por meio de uma topologia de fibra menos malhada, uso de menos comprimentos de onda por enlace e um ROADM mais barato quando comparado à opção *B*. Os exemplos de projeto físico *B* e *C* demonstram que é possível se obter, por meio da nova proposta e no mesmo tempo de execução, configurações de rede que apresentam o custo reduzido a quase metade e apresentando um nível de qualidade semelhante. A Tabela 5.7 também apresenta algumas medidas que resumem a topologia de fibra, como densidade de enlaces,

comprimento médio dos caminhos mínimos e grau médio. Perceba que há um compromisso entre o desempenho da rede e uma métrica como *APL*, pois se os caminhos ópticos percorrem menos nós para chegar ao destino, o sinal é menos degradado. Contudo, é possível obter valores interessantes de *APL* por meio da ligação racional dos enlaces, ao invés de simplesmente acrescentar mais enlaces. Este raciocínio é aplicado na comparação das configurações *A* e *B*. Ambos os projetos *A* e *B* apresentam *APL* inferior a 2, mas o projeto *A* possui um custo inferior e a *PB* é 3 ordens de grandeza menor.

Tabela 5.7: Informações detalhadas de algumas redes que estão destacadas na Figura 5.13.

Rede	<i>PB</i>	Custo (u.m.)	<i>W</i>	δ (dB)	Densidade de enlaces	<i>APL</i>	Grau médio
A	$1,00 \cdot 10^{-5}$	4882,70	37	40	0,25	1,94	3,29
B	$1,66 \cdot 10^{-2}$	6669,93	36	40	0,34	1,77	4,43
C	$2,00 \cdot 10^{-2}$	3470,14	34	35	0,21	2,33	2,71
D	0,84	521,11	4	30	0,14	3,44	1,86

Uma análise detalhada de várias execuções da abordagem evolucionária baseada em CMS permite uma melhor compreensão acerca da utilidade de cada modelo dentro da cascata. A Tabela 5.8 exhibe o tempo médio, a frequência de ocorrência e o percentual de uso de um processo evolucionário completo. É possível observar que o modelo 1 é o mais rápido e é o mais usado também. De fato, a frequência de uso esperada para o modelo 1 seria de 90% se uma quantidade infinita de execuções for considerada. Contudo, o percentual do tempo total de execução é insignificante para o modelo 1. Por outro lado, o modelo 2 é usado em torno de 10% das avaliações e é necessário 1,4 segundos para uma avaliação neste modelo. O modelo 2 corresponde ao gargalo na cascata, pois ele utiliza em torno de 65% do tempo total. Contudo, o ganho líquido no uso da cascata é muito positivo, pois apesar dos modelos 3 até 5 serem computacionalmente custosos, eles estão relacionados com situações mais raras no fluxo evolucionário.

Tabela 5.8: Um resumo do uso de cada modelo substituto em um fluxo completo de execução da nova proposta evolucionária baseada em CMS para a NSFNet.

Modelo	Tempo médio para avaliação (ms)	Frequência de uso	Percentual do tempo total
1	5	0,8878	0,0187
2	1422	0,1083	0,6465
3	9811	0,0034	0,1399
4	88153	0,0003	0,1240
5	856815	0,0001	0,0709

Resultados para ARNES, FLTG e MEDIANET

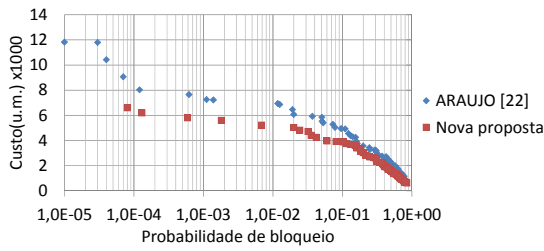
A Tabela 5.9 apresenta um resumo das métricas de qualidade de algoritmos multiobjetivos para a nova proposta e para a abordagem tradicional [22], para cada rede e para os cenários de tráfego uniforme e não-uniforme. A Figura 5.14 apresenta frentes de Pareto de exemplo para comparação do projeto de rede usando a nova proposta e usando a abordagem descrita em Araújo *et. al* [22]. O critério de parada dos algoritmos nesta primeira análise foi o tempo necessário para 1000 avaliações de aptidão para a abordagem baseada estritamente em simulações. Mesmo considerando o mesmo ambiente de execução dos experimentos, este tempo variou de acordo com a rede analisada, a saber: 7 minutos para MEDIANET, 11 minutos para ARNES e 24 minutos para FLTG.

Uma visão geral das métricas de desempenho apresentadas na Tabela 5.9 sugere que a nova abordagem permite obter melhores compromissos de custo e desempenho quando comparada com a abordagem tradicional. O indicador *CS* indica que as soluções da nova abordagem dominam entre 93% (para o caso ARNES, uniforme) e 100% (para FLTG e para ARNES, não-uniforme) das soluções obtidas pela abordagem tradicional. Por outro lado, a abordagem tradicional apresenta soluções que dominam apenas 1% das soluções obtidas pela nova abordagem para o caso MEDIANET, não-uniforme. Contudo, para a maioria dos demais casos analisados, as soluções da abordagem tradicional não domina nenhuma solução da nova proposta quando 1000 avaliações de aptidão são consideradas. Na Tabela 5.9, as células destacadas em negrito para *HV* e *MS* correspondem à abordagem que apresenta o maior valor numérico para estas métricas. É possível perceber que a nova abordagem apresenta maior valor de *HV* em todos os casos analisados e a melhoria deste indicador variou de acordo com a rede analisada, ou seja: há uma melhoria de *HV* de 1% para MEDIANET e em acima de 10% para FLTG e para ARNES. O indicador *MS* apresentou valor maior para metade dos casos considerados. O indicador *MS* deve ser analisado com cautela para frentes de Pareto que ainda não convergiram em direção ao Pareto ótimo, pois as soluções com CAPEX máximo e baixa PB aumentam o valor de *MS*, mas na prática não são configurações de redes vantajosas.

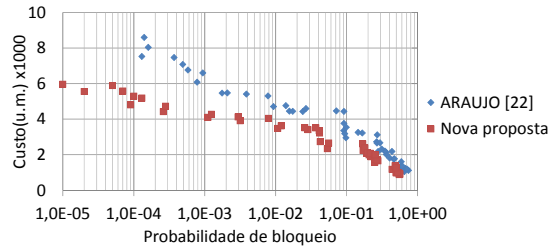
Uma análise panorâmica da Figura 5.14 revela que a nova abordagem é mais efetiva em todos os casos, mas é especialmente mais vantajosa para o projeto de redes com maior quantidade de nós (ARNES e FLTG) e para o caso não-uniforme. É possível perceber, com base na Figura 5.14(a) e na Figura 5.14(b), que a rede MEDIANET é a mais fácil de ser projetada e para a região de baixo custo e elevado PB, as soluções fornecidas pelos dois métodos estão mais próximas. Contudo, esta não é uma região com soluções candidatas para a implantação de redes ópticas reais pois a PB é inactível. Na região de interesse ($PB < 0,10$) a nova abordagem apresenta uma redução de custo de 25% para o caso uniforme (Figura 5.14(a)) e de 30% para o caso não-uniforme (Figura 5.14(b)). Para o projeto de redes ARNES, a nova proposta oferece uma redução de custo de 25% para o caso uniforme

Tabela 5.9: Média e desvio padrão dos indicadores de qualidade, para tráfego uniforme e não-uniforme e para redes FLTG, ARNES e MEDIANET.

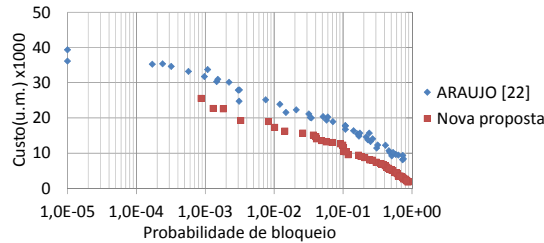
Cenário	Tipo de Tráfego	CS(A, B)		HV		MS	
		CSM, NS	NS, CSM	EA-NS	EA-CSM	EA-NS	EA-CSM
FLTG	Uniforme	0,9750	0,0000	0,8425	0,9288	0,7427	0,7857
		0,0000	0,0000	0,0074	0,0008	0,0270	0,0314
ARNES	Uniforme	1,0000	0,0000	0,8625	0,9513	0,6794	0,4908
		0,0000	0,0000	0,0067	0,0037	0,0637	0,0132
MEDIANET	Uniforme	0,9218	0,0018	0,8193	0,9097	0,7051	0,8151
		0,0515	0,0057	0,0118	0,0013	0,0464	0,0201
ARNES	Não-uniforme	1,0000	0,0000	0,8191	0,9452	0,7138	0,5483
		0,0000	0,0000	0,0150	0,0048	0,0511	0,0236
MEDIANET	Não-uniforme	0,9600	0,0040	0,9272	0,9438	0,7090	0,7372
		0,0326	0,0073	0,0035	0,0014	0,0291	0,0077
ARNES	Uniforme	0,9840	0,0120	0,9560	0,9667	0,6188	0,5809
		0,0213	0,0219	0,0047	0,0015	0,0402	0,0478



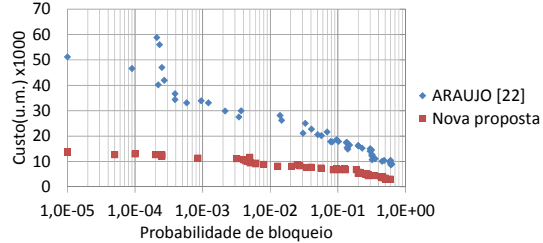
(a) Tráfego uniforme para MEDIANET.



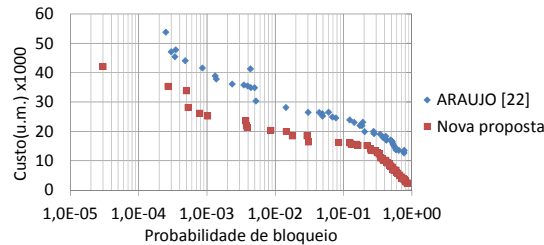
(b) Tráfego não-uniforme para MEDIANET.



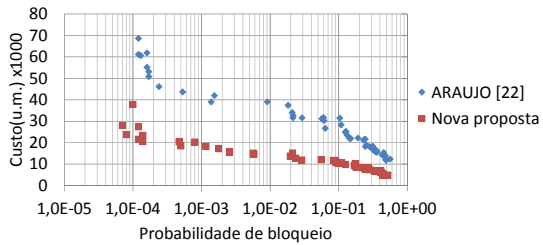
(c) Tráfego uniforme para ARNES.



(d) Tráfego não-uniforme para ARNES.



(e) Tráfego uniforme para FLTG.



(f) Tráfego não-uniforme para FLTG.

Figura 5.14: Frentes de Pareto obtidas após 15 minutos de execução do EA-NS e da nova proposta para redes ARNES, FLTG e MEDIANET.

(Figura 5.14(c)) e de 66% para o caso não-uniforme (Figura 5.14(d)). Para o projeto de redes FLTG, a nova proposta oferece uma redução de custo de 35% para o caso uniforme (Figura 5.14(c)) e de 55% para o caso não-uniforme (Figura 5.14(d)).

Também é interessante observar o nível de desempenho obtido quando o CAPEX é limitado em um projeto de rede. De acordo com a Figura 5.14(a), para o caso uniforme da MEDIANET, é possível obter uma rede que apresenta $PB = 1,3E-4$ e CAPEX de 6000 u.m. e se este mesmo CAPEX for fixado para ARAUJO [22], obtêm-se uma rede com $PB = 2,0E-2$. A mesma análise para o caso ARNES uniforme (Figura 5.14(c)) e para um CAPEX de 22000 u.m., revela que é possível obter uma rede que apresenta $PB = 1,3E-3$ ao invés de outra que apresenta $PB = 1,2E-2$. Para o caso FLTG uniforme (Figura 5.14(e)) e para um CAPEX de 28000 u.m., é possível obter uma rede que apresenta $PB = 5,3E-4$ ao invés de outra que apresenta $PB = 1,4E-2$.

A outra análise de interesse diz respeito à avaliação do tempo necessário para que os algoritmos alcancem um nível determinado de convergência nas frentes de Pareto. Conforme mencionado, foi usado um limiar de variação do indicador *hypervolume* como critério de parada ($\Delta HV < 10^{-6}$). Para o projeto de redes MEDIANET, a nova abordagem alcança a convergência das soluções em média em 13 minutos e a abordagem tradicional requer em média 60 minutos. Para o projeto de redes ARNES, a nova proposta alcança a convergência das soluções em média em 105 minutos e a abordagem tradicional requer em média 285 minutos. Para o projeto de redes FLTG, a nova abordagem alcança a convergência das soluções em média em 291 minutos e a abordagem tradicional requer em média 680 minutos. Portanto, a proposta de planejamento de redes desta Tese oferece uma redução de tempo de 78%, 63% e 57% para os casos MEDIANET, ARNES e FLTG, respectivamente.

5.5 Considerações Finais sobre a Nova Abordagem para Planejamento de Redes

O planejamento de redes ópticas é um problema computacionalmente custoso e geralmente tratado por meio de abordagens heurísticas e meta-heurísticas. Contudo, mesmo as abordagens meta-heurísticas disponíveis na literatura exigem um tempo de resposta que pode ser impraticável dependendo das premissas do projeto. Nesta Tese foi proposto um novo método de planejamento de redes ópticas baseado em conceitos de Ciência das Redes e em Algoritmos Evolucionários. A abordagem proposta possui como dois principais pilares: o uso de procedimentos específicos para geração da topologia física das redes e a proposição de modelos substitutos para avaliação da aptidão de soluções candidatas em processos evolutivos.

De acordo com os resultados apresentados neste capítulo, é possível perceber que a nova abordagem de planejamento oferece ganhos substanciais para os interessados em planejamento de redes ópticas, seja pela redução do tempo exigido para o planejamento ou na obtenção

de soluções de melhor qualidade. Quando um projeto “instantâneo” é requerido, a análise do operador de inicialização revela que o método proposto oferece uma redução de custo de até 69% (caso não-uniforme para ARNES). Quando um tempo máximo é fixado para execução da ferramenta de planejamento, ou seja, a análise de execução do algoritmo após 1000 avaliações de aptidão, revela que a nova proposta oferece uma redução de custo de 25% (caso uniforme para MEDIANET) até 66% (caso não-uniforme para ARNES). Finalmente, se o projeto da rede óptica possui como premissa o fornecimento do melhor conjunto de soluções passível de ser obtido pela meta-heurística, a nova abordagem oferece uma redução de tempo de processamento que varia de 55% (projeto da rede FLTG) até 88% (projeto da rede NSFNet). É importante mencionar que diversos cenários foram analisados no que diz respeito às premissas do projeto da rede, incluindo a variação no número de nós, cobertura geográfica, tipo de tráfego e efeitos da camada física considerados. Portanto, embora os benefícios da nova abordagem sejam quantitativamente diferentes para os casos individuais, na avaliação geral, a nova abordagem oferece vantagens expressivas sobre outras encontradas na literatura.

A proposição da nova abordagem de planejamento e um estudo de caso sobre o projeto da rede NSFNet foi publicada no periódico “*Engineering Applications of Artificial Intelligence*” [130], em 2015.

Capítulo 6

Conclusões

Nesta Tese são propostas diversas abordagens para lidar com o problema de planejamento físico de redes ópticas. O planejamento de redes é tratado nesta Tese por meio de uma abordagem multiobjetiva que considera o custo de implantação da rede (CAPEX) e o desempenho da rede óptica, quantificado por meio da probabilidade de bloqueio. As variáveis de decisão usadas na modelagem de planejamento da Tese incluem a especificação da topologia física e a especificação de dispositivos ópticos.

O problema considerado nesta Tese é classificado como um problema NP-difícil, pois não há algoritmos que forneçam soluções ótimas em um tempo polinomial. Portanto, os trabalhos disponíveis na literatura sobre planejamento físico de redes ópticas fizeram uso de algoritmos heurísticos e meta-heurísticos. Contudo, mesmo considerando o uso de técnicas aproximadas para obtenção de soluções subótimas em um menor tempo, as propostas disponíveis na literatura ainda não oferecem um tempo de planejamento compatível com o uso prático em projetos de engenharia reais. As propostas recentes e que fazem uso de algoritmos evolucionários podem exigir vários dias para finalizar o projeto de uma rede óptica, dependendo da modelagem da camada óptica e das dimensões da rede.

A principal hipótese levantada nesta Tese é a de que os avanços recentes em Ciência das Redes podem contribuir no desenvolvimento de abordagens heurísticas e meta-heurísticas mais sofisticadas para o planejamento de redes ópticas. No Capítulo 3 é feito um levantamento sobre métricas topológicas e modelos de rede que podem ser aplicados no planejamento de redes. Além disso, no Capítulo 3 são propostas métricas próprias que podem ser usadas para quantificar de forma mais precisa diversos aspectos importantes da topologia física e da dinâmica das redes ópticas.

Abordagens de planejamento baseadas em algoritmos evolutivos exigem uma quantificação da qualidade das soluções obtidas para guiar o processo de busca. As abordagens tradicionais fizeram uso de simulações de rede para avaliação do desempenho de uma solução candidata. Contudo as simulações de rede são computacionalmente custosas, especialmente quando as

penalidades da camada física são consideradas para aferição da Qualidade de Transmissão (QoT). No Capítulo 4 é feito um levantamento sobre simuladores de rede que consideram os efeitos da camada física e é proposta uma nova abordagem para estimativa de probabilidade de bloqueio que faz uso combinado das métricas topológicas apresentadas no Capítulo 3 e de redes neurais artificiais (RNAs). A abordagem proposta para estimativa de PB apresenta erros de estimativa na mesma ordem de grandeza dos erros apresentados por simuladores de rede e oferecem uma diminuição no tempo de processamento de até 72 vezes. Foram propostas arquiteturas de RNAs distintas de acordo com o tratamento dado às penalidades da camada física. Portanto, há um compromisso entre precisão nas estimativas e tempo computacional, de modo que o *speedup* obtido varia de acordo com a modelagem adotada para o problema. Além disso, o benefício obtido pelo modelo substituto para avaliação de PB das redes também é ligeiramente diferente quando as demandas de projeto possuem diferentes números de nós e distribuições geográficas. No Capítulo 4 também é feita uma investigação sobre a aplicabilidade da solução proposta para redes com diferentes números de nós e com diferentes coberturas geográficas.

O uso de modelos substitutos para avaliação da aptidão de soluções em algoritmos evolucionários corresponde a uma área de investigação recente. Os modelos substitutos apresentam um compromisso entre fidelidade das análises e tempo de resposta para a avaliação da solução. Caso nenhum controle seja incorporado nos algoritmos evolucionários quando um modelo substituto é usado, o processo evolutivo poderá convergir para falsos pontos ótimos decorrentes de avaliações falhas por parte do modelo substituto. Considerando que a principal motivação do método alternativo proposto no Capítulo 4 é o uso do método para apoiar o planejamento de redes em algoritmos evolutivos, o primeiro passo nesta direção é propor um mecanismo de controle funcional para uso do método. No Capítulo 5 é proposto um novo método de controle de modelos substitutos para algoritmos evolutivos baseado na avaliação da correlação entre o custo da rede e a faixa de desempenho esperado, de acordo com um determinado estágio do processo evolutivo. O método proposto evita que o processo evolucionário convirja para pontos do espaço de busca que foram erroneamente avaliados pelo modelo substituto, mas permite o aproveitamento das avaliações que são consideradas confiáveis. Além disso, o Capítulo 5 propõe novos mecanismos para inicialização de soluções em algoritmos evolucionários inspirados em algoritmos de geração de grafos da Ciência das Redes. Os modelos de geração de redes foram adaptados para considerar aspectos importantes no planejamento de redes ópticas, como a demanda de tráfego da rede. Com base nos estudos realizados foi possível identificar que a diminuição no tempo de avaliação proporcionada pelo método proposto no Capítulo 4 e o novo mecanismo de geração de redes proposto no Capítulo 5 oferece uma ferramenta de planejamento de redes ópticas que exige um tempo de processamento até 88% inferior ao exigido por outras abordagens encontradas na literatura.

Com base nos resultados obtidos nesta Tese, percebe-se que é possível incorporar as es-

estratégias propostas para construção de uma ferramenta de planejamento de redes ópticas, obtendo soluções com um melhor compromisso entre custo e desempenho da rede e em um menor tempo de resposta, adequado para as exigências de prazo que são frequentemente encontradas em projetos de engenharia. Além disso, apesar das propostas apresentadas no Capítulo 3 e no Capítulo 4 terem sido usadas como componentes da solução de planejamento proposta no Capítulo 5, vale a pena mencionar que estas propostas podem ser usadas de forma independente para o desenvolvimento de uma ferramenta própria de análise de redes, evitando o elevado tempo de processamento que é exigido por ferramentas baseadas em simulações.

Também é importante destacar que os avanços obtidos nesta Tese não esgotam a investigação sobre novas estratégias para planejamento de redes ópticas. Os estudos conduzidos nesta Tese focaram em planejamento físico de redes ópticas transparentes. O planejamento lógico de redes ópticas também é um problema NP-difícil e possui potencial para ser beneficiado com os avanços obtidos em Ciência das Redes e em Inteligência Computacional. Trabalhos futuros também podem investigar possíveis adaptações da proposta desta Tese para planejamento físico de redes ópticas translúcidas e opacas. Além disso, os estudos desta Tese demonstraram que o desempenho da solução de planejamento possui dependência com o número de nós da rede a ser planejada. Novos estudos podem ser desenvolvidos com o foco em heurísticas para criação de topologias ou em operadores genéticos mais eficientes para combinação de topologias físicas.

Tabela 6.1: Resumo das publicações e contribuições da Tese.

Publicação	Ano	Contribuições	Tipo de publicação
SBrT [78] (publicado)	2013	Proposição das métricas FZC e HVC	Artigo em conferência
IMOC [92] (publicado)	2013	Proposta inicial de um estimador de PB baseado em RNA	Artigo em conferência
CompleNet [83] (publicado)	2014	Proposição da métrica $\mathcal{I}(\mathcal{F})$	Capítulo de livro
RTIC [84] (publicado)	2014	Avaliação de métricas topológicas para análise de redes ópticas brasileiras	Artigo em periódico, Estrato Qualis B5
JOCN [85] (publicado)	2015	Metodologia para estimador de PB Proposição da métrica CR Tratamento de carga de tráfego na RNA	Artigo em periódico, Qualis A1, Fator de impacto 1,547
EAAI [130] (publicado)	2015	Inicializador heurístico Mecanismo de controle para combinação de modelos substitutos	Artigo em periódico, Qualis A1, Fator de impacto 1,962
ICTON [100] (publicado)	2015	Avaliação do estimador em redes diferentes Novas métricas para uso na RNA	Artigo em conferência
IMOC [101] (aceito)	2015	Proposta de estimador baseado em RNA que não considera efeitos da camada física Análise do impacto do número de redes no treinamento da RNA	Artigo em conferência

6.1 Resumo Quantitativo das Contribuições da Tese

Os trabalhos realizados nesta Tese produziram três artigos publicados em periódicos [84, 85, 130], um capítulo de livro [83], três artigos apresentados em conferências [78, 92, 100] e um artigo aceito para conferência [101]. A Tabela 6.1 relaciona as contribuições realizadas com as publicações em capítulos de livros, periódicos e conferências.

Além disso, as metodologias e algoritmos propostos nesta Tese foram e estão sendo aplicados e analisados em outros trabalhos acadêmicos, a saber: um trabalho de conclusão de curso de graduação concluído [131] e uma dissertação de mestrado em processo de finalização [132].

6.2 Sugestões para Trabalhos Futuros

No início deste capítulo foram fornecidas algumas limitações das abordagens propostas nesta Tese e sugestões para novos estudos. Nesta seção são apresentadas sugestões mais detalhadas para continuidade desta Tese em trabalhos futuros. As sugestões de investigação estão classificadas de acordo com a complexidade do estudo de modo a permitir um direcionamento para futuros trabalhos de graduação e/ou pós-graduação.

Extensões diretas desta Tese e de menor complexidade incluem:

1. Estudo de outras técnicas de aprendizagem para estimativa de PB usando as métricas topológicas consideradas na RNA. Exemplos de técnicas a considerar: *Extreme Learning Machine* (ELM) e *Reservoir Computing* [94];
2. Realização de uma análise sobre o impacto da substituição do SIMTON por outro simulador de rede para ser usado como referência para o motor de aprendizagem;
3. Realização de estudos sobre outros modelos de geração de grafos disponíveis na literatura para serem usados como modelos de topologia física de redes ópticas, com foco em modelos que levam em consideração a disposição geográfica dos nós da rede. São exemplos de modelos geométricos a serem investigados: grafos de Gabriel [133] e de Waxman [134];
4. Realização de estudo de caso sobre o planejamento de redes usando modelos de geração de tráfego baseado em características socio-geográficas [135, 136].

A lista a seguir fornece sugestões de possíveis continuações desta Tese em trabalhos futuros de maior complexidade:

1. Estudo e proposição de operadores genéticos topológicos para melhorar a convergência e diversidade de soluções fornecidas pelo algoritmo de planejamento;

2. Proposição de novos modelos que levem em consideração a cobertura geográfica da rede e as demandas de tráfego (modelos geométricos para geração de grafos);
3. Avaliação de técnicas de compressão de grafos [137] para redução do custo computacional no cálculo de algumas das métricas usadas para estimativa de PB usando RNA;
4. Proposição de novos algoritmos de RWA inspirados em métricas topológicas derivadas da Ciência das Redes que forneçam uma medida da disponibilidade de comprimentos de onda para serem usadas como função de custo;
5. Proposição de modelos alternativos para estimativa de PB que possam ser aplicados em redes translúcidas e opacas;
6. Proposição de modelos alternativos para estimativa de PB que possam ser aplicados em redes ópticas elásticas;
7. Proposição de um mecanismo de controle de modelos substitutos inspirado na abordagem proposta nesta Tese, mas que possa ser aplicado em problema com três ou mais objetivos conflitantes.

Referências

- [1] FREEMAN, L. C. Centrality in social networks conceptual clarification. **Social Networks**, v. 1(79), p. 215–239, 1979.
- [2] RAMASWAMI, R.; SIVARAJAN, K. N. **Optical Networks: A Practical Perspective**. 3^a. [S.l.]: Morgan Kaufmann, 2010.
- [3] SALEH, A. A. M.; SIMMONS, J. M. Technology and architecture to enable the explosive growth of the Internet. **IEEE Communications Magazine**, v. 49, n. 1, p. 126–132, 2011.
- [4] CHIU, A. et al. Network design and architectures for highly dynamic next-generation IP-over-optical long distance networks. **Journal of Lightwave Technology**, v. 27, n. 12, p. 1878–1890, 2009.
- [5] SINKIN, O. V.; HOLZLÖHNER, R.; ZWECK, J.; MENYUK, C. R. Optimization of the Split-Step Fourier Method in Modeling Optical-Fiber Communications Systems. **Journal of Lightwave Technology**, OSA, v. 21, n. 1, p. 61, 2003.
- [6] AGRAWAL, G. P. **Nonlinear Fiber Optics**. 5^a. [S.l.]: Academic Press, 2012. 648 p.
- [7] CHANG, X. Network simulations with opnet. In: ACM. **Proceedings of the 31st conference on Winter simulation: Simulation—a bridge to the future-Volume 1**. [S.l.], 1999. p. 307–314.
- [8] RAMAMURTHY, B.; DATTA, D.; FENG, H. X.; HERITAGE, J. P.; MUKHERJEE, B. SIMON: A simulator for optical networks. In: **Photonics East'99**. [S.l.: s.n.], 1999. p. 130–135.
- [9] WEN, B.; BHIDE, N. M.; SHENAI, R. K.; SIVALINGAM, K. M. Optical wavelength division multiplexing (WDM) network simulator (OWns): Architecture and performance studies. **SPIE Opt. Netw. Mag., Special Issue on Simulation, CAD, and Measurement of Optical Networks, vol.2**, v. 5, 2001.

- [10] SOARES, A.; DURÃES, G.; GIOZZA, W.; CUNHA, P. Tonets: Ferramenta para avaliação de desempenho de redes ópticas transparentes. **VII Salão de Ferramentas do Simpósio Brasileiro de Redes de Computadores-SBRC**, 2008.
- [11] CHAVES, D. A. R.; PEREIRA, H. A.; BASTOS-FILHO, C. J. A.; MARTINS-FILHO, J. F. SIMTON: A simulator for transparent optical networks. **Journal of Communication and Information Systems**, v. 25, p. 1–10, 2010.
- [12] FERNANDEZ, T. F.; MEGALA, T.; SREENATH, N. A comparative survey on optical burst switched network simulators. **International Journal of Computer Network and Information Security**, v. 6, p. 35–41, 2014.
- [13] KO, K.; TANG, K.; CHAN, C.; MAN, K.; KWONG, S. Using genetic algorithms to design mesh networks. **Computer**, v. 30, n. 8, p. 56–61, 1997.
- [14] XIAO G.AND LEUNG, Y.; HUNG, K. Two-stage cut saturation algorithm for designing alloptical networks. **IEEE Transactions on Communications**, v. 49, n. 6, p. 1102–1115, 2001.
- [15] BANNISTER, J.; FRATTA, L.; GERLA, M. Topological design of the wavelength-division optical network. **Ninth Annual Joint Conference of the IEEE Computer and Communication Societies**, v. 3, p. 1005–1013, 1990.
- [16] LIU, H.; TOBAGI, F. A. Physical topology design for all-optical networks. **Optical Switching and Networking**, v. 5, n. 4, p. 219–231, 2008.
- [17] MORAIS, R. M.; PAVAN, C.; PINTO, A. N.; REQUEJO, C. Genetic algorithm for the topological design of survivable optical transport networks. **IEEE/OSA Journal of Optical Communications and Networking**, v. 3, n. 1, p. 17–26, 2011.
- [18] XIN, Y.; ROUSKAS, G. N.; PERROS, H. G. On the Physical and Logical Topology Design of Large-scale Optical Networks. **Journal of Lightwave Technology**, v. 21, p. 904–915, 2003.
- [19] GUAN, C. K.; CHAN, V. Efficient physical topologies for regular WDM networks. In: OPTICAL SOCIETY OF AMERICA. **Optical Fiber Communication Conference**. [S.l.], 2004. p. TuH1.
- [20] SRINIVASAN, A.; PARTHIBAN, R. A heuristic algorithm for topological optimization of all-optical networks. In: IEEE. **2009 WRI World Congress on Computer Science and Information Engineering**. [S.l.], 2009. v. 1, p. 288–292.

- [21] CHAVES, D. A. R. **Projeto de Redes Ópticas de Alta Capacidade Utilizando Técnicas de Otimização Bioinspiradas**. Tese (Doutorado) — Universidade Federal de Pernambuco, 2012.
- [22] ARAÚJO, D. R. B.; BASTOS-FILHO, C. J. A.; BARBOZA, E. A.; CHAVES, D. A. R.; MARTINS-FILHO, J. F. An efficient multi-objective evolutionary optimizer to design all-optical networks considering physical impairments and CAPEX. In: **11th International Conference on Intelligent Systems Design and Applications (ISDA)**. [S.l.: s.n.], 2011. p. 76–81.
- [23] O'MAHONY, M. J.; POLITI, C.; KLONIDIS, D.; NEJABATI, R.; SIMEONIDOU, D. Future Optical Networks. **IEEE/OSA Journal of Lightwave Technology**, v. 24, p. 4684–4696, 2006.
- [24] JINNO, M. et al. Spectrum-efficient and scalable elastic optical path network: architecture, benefits, and enabling technologies. **Communications Magazine, IEEE**, v. 47, n. 11, p. 66–73, November 2009.
- [25] HUELSESMANN, R.; GUNKEL, M.; MEUSBURGER, C.; SCHUPKE, D. A. Cost Modeling and Evaluation of Capital Expenditures in Optical Multilayer Networks. **Journal of Optical Networking**, v. 7, n. 9, p. 814–833, 2008.
- [26] KILPER, D. et al. Optical performance monitoring. **Journal of Lightwave Technology**, v. 22, n. 1, p. 294–304, 2004.
- [27] PAN, Z.; YU, C.; WILLNER, A. E. Optical performance monitoring for the next generation optical communication networks. **Optical Fiber Technology**, v. 16, n. 1, p. 20 – 45, 2010.
- [28] KULKARNI, P.; TZANAKAKI, A.; MACHUKA, C. M.; TOMKOS, I. Benefits of q-factor based routing in WDM metro networks. **Proceedings of ECOC 2005**, IET, v. 4, p. 981–982, 2005.
- [29] MARKIDIS, G.; SYGLETOS, S.; TZANAKAKI, A.; TOMKOS, I. Impairment-constraint-based routing in ultralong-haul optical networks with 2r regeneration. **IEEE Photonics Technology Letters**, v. 19, n. 5/8, p. 420–422, 2007.
- [30] PEREIRA, H. A.; CHAVES, D. A. R.; BASTOS-FILHO, C. A.; MARTINS-FILHO, J. F. OSNR Model to Consider Physical Layer Impairments in Transparent Optical Networks. **Photonic Network Communications**, v. 18, p. 137–149, 2009.
- [31] ZHU, K.; MUKHERJEE, B. Traffic grooming in an optical wdm mesh network. **IEEE Journal on Selected Areas in Communications**, IEEE, v. 20, n. 1, p. 122–133, 2002.

- [32] BANERJEE, N.; KUMAR, R. Multiobjective Network Design for Realistic Traffic Models. **Genetic and Evolutionary Computation (GECCO)**, p. 1904–1910, 2007.
- [33] BARRY, R. A.; HUMBLET, P. A. Models of blocking probability in all-optical networks with and without wavelength changers. **IEEE Journal on Selected Areas in Communications**, v. 14, n. 5, p. 858–867, 1996.
- [34] ALMEIDA-JR, R. C.; CAMPELO, D. R.; WALDMAN, H.; GUILD, K. Accounting for link load correlation in the estimation of blocking probabilities in arbitrary network topologies. **IEEE Communications Letters**, v. 11, p. 625–627, 2007.
- [35] HE, J.; BRANDT-PEARCE, M.; SUBRAMANIAM, S. Analysis of blocking probability for first-fit wavelength assignment in transmission-impaired optical networks. **IEEE/OSA Journal of Optical Communication and Networking**, v. 3, p. 411–425, 2011.
- [36] ABRAMOV, V.; LI, S.; WANG, M.; WONG, E.; ZUKERMAN, M. Computation of blocking probability for large circuit switched networks. **Communications Letters**, v. 16, p. 1892–1895, 2012.
- [37] MCCANNE, S.; FLOYD, S.; FALL, K.; VARADHAN, K. **Network simulator ns-2**. 1997.
- [38] JARGON, J.; WU, X.; WILLNER, A. Optical performance monitoring using artificial neural networks trained with eye-diagram parameters. **IEEE Photonics Technology Letters**, v. 21, n. 1, p. 54–56, 2009.
- [39] SHEN, T.; MENG, K.; LAU, A. P. T.; DONG, Z.-Y. Optical performance monitoring using artificial neural network trained with asynchronous amplitude histograms. **IEEE Photonics Technology Letters**, v. 22, n. 22, p. 1665–1667, 2010.
- [40] DUTTA, R.; ROUSKAS, G. N. A survey of virtual topology design algorithms for wavelength routed optical networks. **Optical Networks**, v. 1, p. 73–89, 2000.
- [41] GROVER, W. D.; DOUCETTE, J. Topological design of survivable mesh-based transport networks. **In: Annals of Operations Research**, p. 79–125, 2001.
- [42] DENGIZ, B.; ALTIPARMAK, F.; SMITH, A. E. Local search genetic algorithm for optimal design of reliable networks. **IEEE Transactions on Evolutionary Computation**, IEEE, v. 1, n. 3, p. 179–188, 1997.
- [43] ELBAUM, R.; SIDI, M. Topological design of local-area networks using genetic algorithms. **IEEE/ACM Transactions on Networking (TON)**, IEEE Press, v. 4, n. 5, p. 766–778, 1996.

- [44] KONAK, A.; SMITH, A. E. A hybrid genetic algorithm approach for backbone design of communication networks. In: **IEEE. Proceedings of the 1999 Congress on Evolutionary Computation**. [S.l.], 1999. v. 3.
- [45] ALTIPARMAK, F.; DENGIZ, B.; SMITH, A. E. Optimal design of reliable computer networks: A comparison of metaheuristics. **Journal of Heuristics**, Springer, v. 9, n. 6, p. 471–487, 2003.
- [46] GERLA, M.; KLEINROCK, L. On the topological design of distributed computer networks. **IEEE Transactions on Communications**, v. 25, p. 48–60, 1977.
- [47] BOORSTYN, R.; FRANK, H. Large-scale network topological optimization. **IEEE Transactions on Communications**, v. 25, p. 29–47, 1977.
- [48] SAYOUD, H.; TAKAHASHI, K.; VAILLANT, B. Designing communication network topologies using steady-state genetic algorithms. **IEEE Communications Letters**, IEEE, v. 5, n. 3, p. 113–115, 2001.
- [49] CHAVES, D. A. R.; BASTOS-FILHO, C. J. A.; MARTINS-FILHO, J. F. Multiobjective Physical Topology Design of All-Optical Networks Considering QoS and Capex. **Conference on Optical Fiber Communication (OFC)**, p. 1–3, 2010.
- [50] MOSS, M. L.; TOWNSEND, A. M. The Internet backbone and the American metropolis. **The Information Society Journal**, v. 16(1), p. 35–47, 2000.
- [51] TRANOS, E.; GILLESPIE, A. The spatial distribution of Internet backbone networks in Europe: A metropolitan knowledge economy perspective. **European Urban and Regional Studies**, v. 16(4), p. 423–437, 2009.
- [52] TRANOS, E. The topology and the emerging urban geographies of the Internet backbone and aviation networks in Europe: a comparative study. **Environment and Planning**, v. 43(2), p. 378–392, 2011.
- [53] CARDENAS, J. P.; BENITO, R. M.; MOURONTE, M. L.; FELIU, V. The effect of the complex topology on the robustness of Spanish SDH network. In: **Fifth International Conference on Networking and Services (ICNS '09)**. [S.l.: s.n.], 2009. p. 86–90.
- [54] KNIGHT, S.; NGUYEN, H. X.; FALKNER, N.; BOWDEN, R.; ROUGHAN, M. The Internet topology zoo. **IEEE Journal on Selected Areas in Communications**, v. 29, n. 9, p. 1765–1775, 2011.
- [55] DHARMAWEERA, N.; PARTHIBAN, R.; SEKERCIOGLU, Y. A. Multi-constraint physical topology design for all optical networks. In: **18th International Conference on Telecommunications (ICT)**. [S.l.: s.n.], 2011. p. 463–469.

- [56] PAIVA, M. H. M.; CAPOROSSO, G.; SEGATTO, M. E. V. Twin graphs for OTN physical topology design. **Les Cahiers du GERAD**, v. 48, p. 1–12, 2013.
- [57] ÇETINKAYA, E. K.; ALENAZI, M. J.; CHENG, Y.; PECK, A. M.; STERBENZ, J. P. A comparative analysis of geometric graph models for modelling backbone networks. **Optical Switching and Networking**, Elsevier, 2014.
- [58] MANIADAKIS, D.; VAROUTAS, D. Incorporating gabriel graph model for FTTx dimensioning. **Photonic Network Communications**, Springer, v. 29, n. 2, p. 214–226, 2015.
- [59] APPLICATIONS, N. R. C. Committee on Network Science for F. A. **Network Science**. [S.l.]: The National Academies Press, 2005.
- [60] ERDOS, P.; RÉNYI, A. On the evolution of random graphs. **Publication of the Mathematical Institute of the Hungarian Academy of Sciences**, p. 17–61, 1960.
- [61] WATTS, D. J.; STROGATZ, S. H. Collective dynamics of small-world networks. **Nature**, v. 393, p. 440–442, 1998.
- [62] BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. **Science**, v. 286, p. 509–512, 1999.
- [63] LEWIS, T. G. **Network Science - Theory and Applications**. [S.l.]: John Wiley & Sons, 2009.
- [64] ALBERT, R.; JEONG, H.; BARABÁSI, A. L. Error and attack tolerance of complex networks. **Nature**, v. 406(6794), p. 378–382, 2000.
- [65] SHA, Z.; PANCHAL, J. Towards the design of complex evolving networks with high robustness and resilience. **Procedia Computer Science**, v. 16, p. 522–531, 2013.
- [66] BOLLOBÁS, B. **Modern graph theory**. [S.l.]: Springer-Verlag, 1998.
- [67] WU, J.; BARAHONA, M.; TAN, Y.-J.; DENG, H.-Z. Spectral measure of structural robustness in complex networks. **IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans**, v. 41, n. 6, p. 1244 –1252, 2011.
- [68] HOLME, P.; KIM, B. J.; YOON, C. N.; HAN, S. K. Attack vulnerability of complex networks. **Physical Review**, v. 65, n. 5, p. 056–109, 2002.
- [69] DAM, E. R. van; HAEMERS, W. H. Which graphs are determined by their spectrum? **Linear Algebra and its Applications**, v. 373, p. 241–272, 2003.

- [70] WANG, Y.; CHAKRABARTI, D.; WANG, C.; FALOUTSOS, C. Epidemic spreading in real networks: An eigenvalue viewpoint. **Proceedings of the 22nd Symposium in Reliable Distributed Computing**, 2003.
- [71] FIEDLER, M. Algebraic connectivity of graphs. **Czechoslovak Mathematical Journal**, v. 23, p. 298–305, 1973.
- [72] MIEGHEM, P. van. **Graph Spectra for Complex Networks**. [S.l.]: Cambridge University Press, 2011.
- [73] RAO, K. R.; KIM, D. N.; HWANG, J.-J. **Fast Fourier Transform - Algorithms and Applications**. [S.l.]: Springer, 2010. 442 p. (Signals and Communication Technology).
- [74] FOUNDATION, A. S. **Commons Math: The Apache Commons Mathematics Library**. <http://commons.apache.org/proper/commons-math/>, 2014. Último acesso: 2015-06-21.
- [75] TRIOLA, M. F. **Elementary Statistics**. 12th. ed. [S.l.]: Pearson, 2012.
- [76] ISELLA, L. et al. What's in a Crowd? Analysis of Face-to-face Behavioral Networks. **Journal. of Theoretical Biology**, v. 271, n. 1, p. 166–180, 2011.
- [77] JEONG, H.; TOMBOR, B.; ALBERT, R.; OLTVAI, Z. N.; BARABÁSI, A. L. The large-scale organization of metabolic networks. **Nature**, v. 407, p. 651–654, 2000.
- [78] ARAÚJO, D. R. B.; BASTOS-FILHO, C. J. A.; MARTINS-FILHO, J. F. Towards Using DFT to Characterize Complex Networks. In: **XXXI Simpósio Brasileiro de Telecomunicações (SBrT2013)**. [S.l.: s.n.], 2013. p. 1–5.
- [79] READ, K. E. Cultures of the central highlands. **Southwestern J. of Anthropology**, v. 10(1), p. 1–43, 1954.
- [80] ZACHARY, W. W. An information flow model for conflict and fission in small groups. **Journal of Anthropological Research**, v. 33, p. 452–473, 1977.
- [81] MICHALSKI, R.; PALUS, S.; KAZIENKO, P. Matching organizational structure and social network extracted from email communication. **Lecture Notes in Business Information Processing**, v. 87, p. 197–206, 2011.
- [82] GUIMERA, R.; DANON, L.; DIAZ-GUILERA, A.; GIRALT, F.; ARENAS, A. Self-similar community structure in a network of human interactions. **Physical Review E**, v. 68, p. 065–103, 2003.

- [83] ARAÚJO, D. R. B.; BASTOS-FILHO, C. J. A.; MARTINS-FILHO, J. F. Complex networks v. In: _____. [S.l.]: Springer, 2014. (Studies in Computational Intelligence, 1), cap. Using the Entropy of DFT of the Laplacian Eigenvalues to Assess Networks, p. 209–216.
- [84] ARAÚJO, D. R. B.; BASTOS-FILHO, C. J. A.; MARTINS-FILHO, J. F. Métricas de redes complexas para análise de redes Ópticas. **Revista de Tecnologia da Informação e Comunicação**, v. 4, n. 2, p. 1–8, 2014.
- [85] ARAÚJO, D. R. B.; BASTOS-FILHO, C. J. A.; MARTINS-FILHO, J. F. Methodology to Obtain a Fast and Accurate Estimator for Blocking Probability of Optical Networks. **Journal of Optical Communications and Networking**, v. 7, n. 5, p. 380–391, Maio 2015.
- [86] BONANI, L. H.; FONSECA, I. E. Estimating the blocking probability in wavelength-routed optical networks. **Optical Switching and Networking**, v. 10, n. 4, p. 430–438, 2013.
- [87] ARAÚJO, D. R. B. **Comparação de Algoritmos de Otimização Multiobjetivos Aplicados a Projeto de Redes Ópticas**. Dissertação (Mestrado) — Universidade de Pernambuco, 2011.
- [88] TANENBAUM, A. S.; WETHERALL, D. J. **Computer Networks**. 5^a. [S.l.]: Prentice Hall, 2010. 960 p.
- [89] YANG, X.; SHEN, L.; RAMAMURTHY, B. Survivable lightpath provisioning in WDM mesh networks under shared path protection and signal quality constraints. **IEEE/OSA Journal of Lightwave Technology**, v. 23, n. 4, p. 1556–1567, 2005.
- [90] ZULKIFLI, N.; ALMEIDA-JR, R. C.; GUILD, K. M. Efficient resource allocation of heterogeneous services in transparent optical networks. **Journal of Optical Networking**, v. 6, n. 12, p. 1349–1359, 2007.
- [91] ZANG, H.; JUE, J. P.; MUKHERJEE, B. A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks. **Optical Networks Magazine**, v. 1, n. 1, p. 47–60, 2000.
- [92] ARAÚJO, D. R. B.; BASTOS-FILHO, C. J. A.; MARTINS-FILHO, J. F. Using Multi-Layer Perceptron and Complex Network Metrics to Estimate the Performance of Optical Networks. In: **Proceedings of the SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference**. [S.l.: s.n.], 2013. v. 1, p. 1–5.

- [93] HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning**. [S.l.]: Springer, 2008.
- [94] HAYKIN, S. O. **Neural Networks and Learning Machines**. 3^a. [S.l.]: Prentice Hall, 2008. 936 p.
- [95] FLTG Network Map. [Http://www.fltg.com/pop-NetworkMap.html](http://www.fltg.com/pop-NetworkMap.html). Último acesso: 2015-06-21.
- [96] Arnes Infrastructure - basis for providing services. [Http://www.arnes.si/en/infrastructure.html](http://www.arnes.si/en/infrastructure.html). Último acesso: 2015-06-21.
- [97] HSE Medianet. [Https://www.hse-medianet.de](https://www.hse-medianet.de). Último acesso: 2015-06-21.
- [98] PALLOTTINO, S. Shortest-path methods: Complexity, interrelations and new propositions. **Networks**, Wiley Online Library, v. 14, n. 2, p. 257–267, 1984.
- [99] ARAÚJO, D. R. B.; BASTOS-FILHO, C. J. A.; BARBOZA, E. A.; CHAVES, D. A. R.; MARTINS-FILHO, J. F. A Performance Comparison of Multi-Objective Optimization Evolutionary Algorithms for All-Optical Networks Design. **IEEE Symposium on Computational Intelligence in Multicriteria Decision-Making**, 2011.
- [100] ARAÚJO, D. R. B.; BASTOS-FILHO, C. J. A.; MARTINS-FILHO, J. F. Artificial neural networks to estimate blocking probability of transparent optical networks: A robustness study for different networks. In: **Proceedings of the 17th International Conference on Transparent Optical Networks (ICTON)**. [S.l.: s.n.], 2015.
- [101] ARAÚJO, D. R. B.; FILHO, C. J. A. B.; MARTINS-FILHO, J. F. Analyzing surrogate models to assess blocking probability of optical networks. In: **Proceeding of the SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC)**. [S.l.: s.n.], 2015. p. 1–5.
- [102] ZHANGA, R.; KABADIB, S. N.; PUNNEN, A. P. The minimum spanning tree problem with conflict constraints and its variations. **Discrete Optimization**, v. 8(2), p. 191–205, 2011.
- [103] EIBEN, A.; SMITH, J. E. **Introduction to Evolutionary Computing**. [S.l.]: Springer, 2008. 300 p. (Natural Computing Series).
- [104] COELLO-COELLO, C. A.; LAMONT, G. B.; VELDHUIZEN, D. A. V. **Evolutionary Algorithms for Solving Multi-Objective Problems**. 2^a. [S.l.]: Springer, 2007.
- [105] KNOWLES, J. D.; CORNE, D. W. A Comparison of Encodings and Algorithms for Multiobjective Minimum Spanning Tree Problems. In: **Proceedings of the Congress on Evolutionary Computation (CEC-01)**. [S.l.: s.n.], 2001. v. 1, p. 544– 551.

- [106] KUMAR, R.; SINGH, P. K.; CHAKRABARTI, P. P. Multiobjective EA Approach for Improved Quality of Solutions for Spanning Tree Problem. In: **Proceedings of the International Conference on Evolutionary Multi-Criterion Optimization (EMO-05)**. [S.l.: s.n.], 2005. p. 811–825.
- [107] KNOWLES, J. ParEGO: A Hybrid Algorithm With On-Line Landscape Approximation for Expensive Multiobjective Optimization Problems. **IEEE Transactions on Evolutionary Computation**, v. 10(1), p. 50–66, 2006.
- [108] MARTINEZ, S. Z.; COELLO-COELLO, C. Combining surrogate models and local search for dealing with expensive multi-objective optimization problems. In: **2013 IEEE Congress on Evolutionary Computation (CEC)**. [S.l.: s.n.], 2013. p. 2572–2579.
- [109] SCHOTT, J. **Fault Tolerant Design Using Single and Multicriteria Genetic Algorithm Optimization**. Tese (Doutorado) — Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1995.
- [110] VELDHUIZEN, D. A. V.; LAMONT, G. B. Evolutionary computation and convergence to a pareto front. In: **Late Breaking Papers at the Genetic Programming Conference**. [S.l.: s.n.], 1998. p. 221–228.
- [111] KNOWLES, J.; CORNE, D. On Metrics for Comparing Nondominated Sets. **Congress on Evolutionary Computation (CEC)**, v. 1, p. 711–716, 2002.
- [112] ZITZLER, E.; THIELE, L.; LAUMANN, M.; FONSECA, C.; FONSECA, V. Grunert da. Performance Assessment of Multiobjective Optimizers: An Analysis and Review. **IEEE Transactions on Evolutionary Computation**, v. 7(2), p. 117–132, 2003.
- [113] ZITZLER, E. **Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications**. Tese (Doutorado) — ETH Zurich, Switzerland, 1999.
- [114] DEB, K.; PRATAP, A.; AGARWAL, S.; MEYARIVAN, T. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. **IEEE Transactions on Evolutionary Computation**, v. 6(2), p. 182–197, 2002.
- [115] DEB, K.; JAIN, H. An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach, Part I: Solving Problems With Box Constraints. **IEEE Transactions on Evolutionary Computation**, v. 18, n. 4, p. 577–601, 2014.
- [116] JIN, Y.; SENDHOFF, B. A systems approach to evolutionary multiobjective structural optimization and beyond. **IEEE Computational Intelligence Magazine**, IEEE, v. 4, n. 3, p. 62–76, 2009.

- [117] DOUGUET, D. A E-LEA3D: a computational-aided drug design web server. **Nucleic acids research**, Oxford Univ Press, p. 1–7, 2010.
- [118] SCHNEIDER, G.; SCHUCHHARDT, J.; WREDE, P. Artificial neural networks and simulated molecular evolution are potential tools for sequence-oriented protein design. **Computer applications in the biosciences: CABIOS**, Oxford Univ Press, v. 10, n. 6, p. 635–645, 1994.
- [119] RATLE, A. Accelerating the convergence of evolutionary algorithms by fitness landscape approximation. In: SPRINGER. **Parallel Problem Solving from Nature-PPSN V**. [S.l.], 1998. p. 87–96.
- [120] JIN, Y. Surrogate-assisted evolutionary computation: Recent advances and future challenges. **Swarm and Evolutionary Computation**, Elsevier, v. 1, n. 2, p. 61–70, 2011.
- [121] JIN, Y.; SENDHOFF, B. Reducing fitness evaluations using clustering techniques and neural network ensembles. In: SPRINGER. **Genetic and Evolutionary Computation (GECCO)**. [S.l.], 2004. p. 688–699.
- [122] TENNE, Y.; ARMFIELD, S. W. A framework for memetic optimization using variable global and local surrogate models. **Soft Computing**, Springer, v. 13, n. 8-9, p. 781–793, 2009.
- [123] JIN, Y.; OLHOFER, M.; SENDHOFF, B. A framework for evolutionary optimization with approximate fitness functions. **IEEE Transactions on Evolutionary Computation**, IEEE, v. 6, n. 5, p. 481–494, 2002.
- [124] BRANKE, J.; SCHMIDT, C. Faster convergence by means of fitness estimation. **Soft Computing**, Springer, v. 9, n. 1, p. 13–20, 2005.
- [125] SCHMIDT, M. D.; LIPSON, H. Coevolution of fitness predictors. **IEEE Transactions on Evolutionary Computation**, IEEE, v. 12, n. 6, p. 736–749, 2008.
- [126] TENNE, Y.; IZUI, K.; NISHIWAKI, S. Dimensionality-reduction frameworks for computationally expensive problems. In: IEEE. **IEEE Congress on Evolutionary Computation (CEC)**. [S.l.], 2010. p. 1–8.
- [127] MUKHERJEE, B.; BANERJEE, D.; RAMAMURTHY, S.; MUKHERJEE, A. Some principles for designing a wide-area wdm network. **IEEE/ACM Transactions on Networking**, v. 4(5), p. 684–696, 1996.
- [128] KRINGS, G.; CALABRESE, F.; RATTI, C.; BLONDEL, V. D. Urban gravity: a model for inter-city telecommunication flows. **Journal of Statistical Mechanics: Theory and Experiment**, IOP Publishing, v. 2009, n. 07, p. 1–8, 2009.

- [129] MAN, K.; TANG, K.; KWONG, S. Genetic Algorithms: Concepts and Applications (in Engineering Design). **IEEE Transactions on Industrial Electronics**, v. 43-5, p. 519–534, 1996.
- [130] ARAÚJO, D. R. B.; BASTOS-FILHO, C. J. A.; MARTINS-FILHO, J. F. An evolutionary approach with surrogate models and network science concepts to design optical networks. **Engineering Applications of Artificial Intelligence**, v. 43, n. 8, p. 67–80, 2015.
- [131] SILVA, J. C. **Avaliação de Robustez a Variações de Carga de Um Estimador de Probabilidade de Bloqueio para Redes Ópticas Baseado em Redes Neurais**. Monografia (Conclusão de Curso de Graduação) — Universidade de Pernambuco, 2014.
- [132] LIMA, N. J. S. **Algoritmos de sobrevivência para redes ópticas**. Dissertação (Mestrado em Andamento) — Universidade de Pernambuco, 2015.
- [133] MATULA, D. W.; SOKAL, R. R. Properties of gabriel graphs relevant to geographic variation research and the clustering of points in the plane. **Geographical Analysis**, v. 12, n. 3, p. 205–222, 1980.
- [134] WAXMAN, B. M. Routing of multipoint connections. **IEEE Journal on Selected Areas in Communications**, v. 6, n. 9, p. 1617–1622, 1988.
- [135] DWIVEDI, A.; WAGNER, R. Traffic model for USA long-distance optical network. In: **Optical Fiber Communication Conference**. [S.l.: s.n.], 2000. p. 156–158.
- [136] ROUGHHAN, M. et al. Experience in measuring backbone traffic variability: models, metrics, measurements and meaning. In: **ACM SIGCOMM Internet Measurement Workshop**. [S.l.: s.n.], 2002. p. 91–92.
- [137] ALAMSYAH, A.; PERANGINANGIN, Y.; RAHARDJO, B.; MUCHTADI-ALAMSYAH, I. et al. Reducing computational complexity of network analysis using graph compression method for brand awareness effort. In: ATLANTIS PRESS. **3rd International Conference on Computation for Science and Technology (ICCST-3)**. [S.l.], 2015.

Apêndice A

Combinações de Variáveis Independentes para Estimativa de PB

Neste apêndice são fornecidas as melhores combinações de variáveis para estimativa de PB das redes MEDIANET, ARNES e FLTG. Nos casos em que mais de uma combinação com o mesmo número de variáveis apresenta uma diferença nos erros de estimativa de menos de 5%, é apresentada mais de uma linha para a mesma quantidade de variáveis.

Tabela A.1: Melhores modelos de regressão com p variáveis para estimativa de PB para o problema Tipo A e para a rede MEDIANET.

p	Camada de entrada	\overline{MSE}	σ_{MSE}	$\Delta\overline{MSE}$
2	$W, \overline{\Delta OSNR}$	4,86E-3	1,81E-4	-
	W, CR	4,93E-3	1,44E-5	-
3	$W, L, \overline{\Delta OSNR}$	1,80E-3	2,11E-6	0,63
4	$W, \delta, L, \overline{\Delta OSNR}$	7,00E-4	1,80E-5	0,61
5	$W, \delta, CR, L, \overline{\Delta OSNR}$	4,06E-4	1,49E-5	0,42
6	$W, \delta, CR, L, q_{km}, \overline{\Delta OSNR}$	3,34E-4	4,19E-6	0,18
	$W, \delta, q, CR, L, \overline{\Delta OSNR}$	3,48E-4	5,50E-6	0,14
	$W, \delta, \bar{c}_{km}, CR, L, \overline{\Delta OSNR}$	3,49E-4	1,90E-5	0,14
7	$W, \delta, q, \bar{c}_{km}, CR, L, \overline{\Delta OSNR}$	3,13E-4	2,20E-6	0,06
8	$W, \delta, CC, q, \bar{c}_{km}, CR, L, \overline{\Delta OSNR}$	3,05E-4	6,02E-6	0,03

Tabela A.2: Melhores modelos de regressão com p variáveis para estimativa de PB para o problema Tipo B e para a rede MEDIANET.

p	Camada de entrada	\overline{MSE}	σ_{MSE}	$\Delta\overline{MSE}$
2	W, \bar{c}	3,23E-3	6,37E-6	-
3	W, \bar{c}, L	4,78E-4	4,66E-6	0,85
4	W, q, \bar{c}, L	3,52E-4	5,50E-6	0,29
5	W, CC, q, \bar{c}, L	3,09E-4	4,30E-6	0,09
	W, q, \bar{c}, CR, L	3,16E-4	2,65E-6	0,06
6	W, CC, q, \bar{c}, CR, L	2,99E-4	1,97E-6	0,03

Tabela A.3: Melhores modelos de regressão com p variáveis para estimativa de PB para o problema Tipo A e para a rede ARNES.

p	Camada de entrada	\overline{MSE}	σ_{MSE}	$\Delta\overline{MSE}$
2	W, q_{km}	9,53E-3	4,58E-4	-
3	W, L, q_{km}	4,60E-3	9,11E-6	0,52
4	W, δ, L, q_{km}	1,05E-3	6,18E-5	0,77
5	W, δ, q, L, q_{km}	7,27E-4	9,14E-5	0,31
	$W, \delta, L, q_{km}, \overline{\Delta OSNR}$	7,31E-4	1,94E-5	0,30
6	$W, \delta, q, CR, L, q_{km}$	4,18E-4	1,64E-5	0,42
7	$W, \delta, q, \bar{c}_{km}, CR, L, q_{km}$	2,97E-4	8,65E-6	0,29
	$W, \delta, q, I(\hat{\mathcal{F}}), CR, L, q_{km}$	3,42E-4	1,52E-5	0,18
8	$W, \delta, q, I(\hat{\mathcal{F}}), CR, L, q_{km}, \sigma_{PL}$	2,97E-4	1,17E-5	0,00

Tabela A.4: Melhores modelos de regressão com p variáveis para estimativa de PB para o problema Tipo B e para a rede ARNES.

p	Camada de entrada	\overline{MSE}	σ_{MSE}	$\Delta\overline{MSE}$
2	W, \bar{c}	1,99E-3	2,16E-5	-
3	W, \bar{c}, L	4,31E-4	3,91E-6	0,78
4	$W, \bar{c}, I(\hat{\mathcal{F}}), L$	3,92E-4	3,90E-6	0,08
	W, \bar{c}, CR, L	4,18E-4	3,33E-6	0,03
5	$W, q, \bar{c}, I(\hat{\mathcal{F}}), L$	3,36E-4	4,31E-6	0,14
6	$W, CC, q, \bar{c}, I(\hat{\mathcal{F}}), L$	3,23E-4	3,17E-6	0,04

Tabela A.5: Melhores modelos de regressão com p variáveis para estimativa de PB para o problema Tipo A e para a rede FLTG.

p	Camada de entrada	\overline{MSE}	σ_{MSE}	$\overline{\Delta MSE}$
2	W, q	1,29E-2	2,00E-4	-
	W, CR	1,30E-2	4,39E-5	-
	$W, \overline{\Delta OSNR}$	1,34E-2	7,96E-5	-
	W, q_{km}	1,35E-2	5,15E-5	-
3	W, δ, q	5,19E-3	2,66E-4	0,60
4	W, δ, q, L	1,45E-3	1,30E-4	0,72
5	W, δ, q, CR, L	7,79E-4	3,07E-5	0,46
6	$W, \delta, q, CR, L, \overline{\Delta OSNR}$	7,21E-4	4,56E-5	0,07
	$W, \delta, q, CR, L, \sigma_{\Delta OSNR}$	7,53E-4	1,08E-4	0,03
	W, δ, CC, q, CR, L	7,56E-4	6,92E-5	0,03
7	$W, \delta, q, \bar{c}_{km}, CR, L, \overline{\Delta OSNR}$	6,23E-4	4,33E-5	0,27
8	$W, \delta, q, \bar{c}_{km}, CR, L, \overline{\Delta OSNR}, \sigma_{\Delta OSNR}$	5,43E-4	1,96E-5	0,13
9	$W, \delta, q, \bar{c}_{km}, CR, L, \overline{\Delta OSNR}, \sigma_{PL}, \sigma_{\Delta OSNR}$	5,25E-4	6,3E-5	0,03

Tabela A.6: Melhores modelos de regressão com p variáveis para estimativa de PB para o problema Tipo B e para a rede FLTG.

p	Camada de entrada	\overline{MSE}	σ_{MSE}	$\overline{\Delta MSE}$
2	W, \bar{c}	1,76E-3	1,98E-5	-
3	W, \bar{c}, L	6,54E-4	4,99E-6	0,63
4	W, q, \bar{c}, L	5,57E-4	4,73E-6	0,15
5	$W, q, \bar{c}, I(\hat{\mathcal{F}}), L$	5,30E-4	9,82E-6	0,05
6	$W, CC, q, \bar{c}, I(\hat{\mathcal{F}}), L$	5,28E-4	5,61E-6	0,00

Apêndice B

Matriz de Tráfego das Redes Analisadas

Para cada rede analisada nesta Tese foi considerada uma matriz de tráfego compatível com o tráfego das redes originais.

Para a rede NSFNet de 14 nós, a matriz de tráfego foi obtida a partir do estudo [127], no qual foi divulgado o tráfego coletado nesta rede na época em que o estudo foi realizado. Uma versão bidirecional desta matriz foi construída e está ilustrada na Tabela B.1.

Para as redes MEDIANET, ARNES e FTLG foram criadas matrizes de tráfego usando um modelo gravitacional [128] que considera a população associada à cada nó da rede e a distância física entre os nós. A matriz usada para a rede MEDIANET está ilustrada na Tabela B.2. A matriz usada para a rede ARNES está ilustrada nas Tabelas B.3 e B.3. A matriz usada para a rede FTLG está ilustrada nas Tabelas B.5 e B.6.

Tabela B.1: Matriz de tráfego usada para o cenário de tráfego não-uniforme para a rede NSFNet.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	0,00	6,40	1,03	0,23	0,37	0,02	1,09	1,92	3,48	0,08	0,09	0,69	0,00	0,53
2	6,40	0,00	4,28	2,02	39,32	6,79	12,82	39,82	11,30	25,95	1,53	31,98	2,42	21,14
3	1,03	4,28	0,00	6,82	9,51	1,71	51,16	105,15	18,68	5,13	1,57	0,50	0,00	6,82
4	0,23	2,02	6,82	0,00	2,67	3,06	2,59	5,70	4,00	3,50	2,38	6,32	0,04	6,95
5	0,37	39,32	9,51	2,67	0,00	3,72	16,40	45,22	24,51	20,13	0,35	51,95	4,99	20,45
6	0,02	6,79	1,71	3,06	3,72	0,00	5,26	2,67	3,84	6,68	0,05	1,08	0,03	7,28
7	1,09	12,82	51,16	2,59	16,40	5,26	0,00	105,59	22,44	13,47	0,80	38,45	0,17	19,64
8	1,92	39,82	105,15	5,70	45,22	2,67	105,59	0,00	52,49	45,04	6,41	44,60	2,77	21,32
9	3,48	11,30	18,68	4,00	24,51	3,84	22,44	52,49	0,00	23,23	7,55	78,52	11,67	2,72
10	0,08	25,95	5,13	3,50	20,13	6,68	13,47	45,04	23,23	0,00	6,71	20,01	2,96	0,38
11	0,09	1,53	1,57	2,38	0,35	0,05	0,80	6,41	7,55	6,71	0,00	5,67	0,03	2,25
12	0,69	31,98	0,50	6,32	51,95	1,08	38,45	44,60	78,52	20,01	5,67	0,00	0,51	0,27
13	0,00	2,42	0,00	0,04	4,99	0,03	0,17	2,77	11,67	2,96	0,03	0,51	0,00	5,38
14	0,53	21,14	6,82	6,95	20,45	7,28	19,64	21,32	2,72	0,38	2,25	0,27	5,38	0,00

Tabela B.2: Matriz de tráfego usada para o cenário de tráfego não-uniforme para a rede MEDIANET.

1	0,00	19,11	1,96	1,78	0,47	0,57	2,04	2,08	200,00	34,29	5,98	12,96	10,87	3,37	2,92	19,09	70,92	6,77
2	19,11	0,00	8,81	2,31	0,29	0,36	1,41	1,93	4,54	1,91	1,32	3,83	3,51	3,52	17,84	5,09	1,05	4,10
3	1,96	8,81	0,00	0,96	0,05	0,05	0,15	0,15	0,43	0,20	0,16	0,29	0,42	0,41	2,19	0,77	0,10	1,02
4	1,78	2,31	0,96	0,00	0,12	0,07	0,16	0,12	0,38	0,21	0,19	0,19	0,30	0,21	0,44	0,73	0,08	3,71
5	0,47	0,29	0,05	0,12	0,00	0,09	0,09	0,04	0,11	0,07	0,12	0,03	0,05	0,02	0,04	0,11	0,02	0,40
6	0,57	0,36	0,05	0,07	0,09	0,00	0,42	0,10	0,15	0,11	0,22	0,04	0,05	0,03	0,04	0,10	0,02	0,19
7	2,04	1,41	0,15	0,16	0,09	0,42	0,00	1,19	0,60	0,40	0,67	0,15	0,15	0,08	0,14	0,29	0,09	0,41
8	2,08	1,93	0,15	0,12	0,04	0,10	1,19	0,00	0,65	0,35	0,35	0,19	0,15	0,09	0,17	0,27	0,09	0,29
9	200,00	4,54	0,43	0,38	0,11	0,15	0,60	0,65	0,00	14,39	1,77	2,41	1,63	0,61	0,62	2,75	5,10	1,33
10	34,29	1,91	0,20	0,21	0,07	0,11	0,40	0,35	14,39	0,00	2,19	0,65	0,59	0,23	0,26	1,19	0,85	0,79
11	5,98	1,32	0,16	0,19	0,12	0,22	0,67	0,35	1,77	2,19	0,00	0,26	0,27	0,13	0,17	0,59	0,21	0,71
12	12,96	3,83	0,29	0,19	0,03	0,04	0,15	0,19	2,41	0,65	0,26	0,00	2,56	1,06	0,64	2,03	1,19	0,56
13	10,87	3,51	0,42	0,30	0,05	0,05	0,15	0,15	1,63	0,59	0,27	2,56	0,00	3,10	0,88	13,29	0,72	1,05
14	3,37	3,52	0,41	0,21	0,02	0,03	0,08	0,09	0,61	0,23	0,13	1,06	3,10	0,00	1,37	2,30	0,21	0,56
15	2,92	17,84	2,19	0,44	0,04	0,04	0,14	0,17	0,62	0,26	0,17	0,64	0,88	1,37	0,00	1,23	0,16	0,80
16	19,09	5,09	0,77	0,73	0,11	0,10	0,29	0,27	2,75	1,19	0,59	2,03	13,29	2,30	1,23	0,00	1,02	3,23
17	70,92	1,05	0,10	0,08	0,02	0,02	0,09	0,09	5,10	0,85	0,21	1,19	0,72	0,21	0,16	1,02	0,00	0,30
18	6,77	4,10	1,02	3,71	0,40	0,19	0,41	0,29	1,33	0,79	0,71	0,56	1,05	0,56	0,80	3,23	0,30	0,00

Tabela B.3: Matriz de tráfego usada para o cenário de tráfego não-uniforme para a rede ARNES (continua).

1	0,00	0,72	19,24	2,87	2,09	0,15	2,43	23,68	0,06	0,45	0,03	0,06	0,04	0,40	0,01	0,33	0,09
2	0,72	0,00	2,40	0,31	0,34	0,03	1,99	3,63	0,01	0,13	0,01	0,03	0,02	0,36	0,00	0,07	0,03
3	19,24	2,40	0,00	18,90	3,23	0,28	4,27	30,33	0,11	0,84	0,05	0,24	0,16	2,50	0,02	0,63	0,16
4	2,87	0,31	18,90	0,00	0,30	0,03	0,58	3,14	0,01	0,08	0,01	0,02	0,01	0,18	0,00	0,06	0,02
5	2,09	0,34	3,23	0,30	0,00	3,95	1,55	200,00	0,63	3,45	0,12	0,07	0,05	0,22	0,07	5,86	0,48
6	0,15	0,03	0,28	0,03	3,95	0,00	0,13	7,45	0,16	0,57	0,02	0,01	0,01	0,02	0,01	13,19	0,06
7	2,43	1,99	4,27	0,58	1,55	0,13	0,00	21,09	0,07	0,61	0,05	0,06	0,04	0,36	0,02	0,29	0,14
8	23,68	3,63	30,33	3,14	200,00	7,45	21,09	0,00	2,88	22,35	1,02	0,54	0,36	1,91	0,71	13,92	4,45
9	0,06	0,01	0,11	0,01	0,63	0,16	0,07	2,88	0,00	1,91	0,02	0,00	0,00	0,01	0,01	0,37	0,08
10	0,45	0,13	0,84	0,08	3,45	0,57	0,61	22,35	1,91	0,00	0,43	0,03	0,02	0,08	0,14	1,28	1,94
11	0,03	0,01	0,05	0,01	0,12	0,02	0,05	1,02	0,02	0,43	0,00	0,00	0,00	0,01	0,01	0,04	0,19
12	0,06	0,03	0,24	0,02	0,07	0,01	0,06	0,54	0,00	0,03	0,00	0,00	3,11	0,06	0,00	0,02	0,01
13	0,04	0,02	0,16	0,01	0,05	0,01	0,04	0,36	0,00	0,02	0,00	3,11	0,00	0,04	0,00	0,01	0,00
14	0,40	0,36	2,50	0,18	0,22	0,02	0,36	1,91	0,01	0,08	0,01	0,06	0,04	0,00	0,00	0,05	0,01
15	0,01	0,00	0,02	0,00	0,07	0,01	0,02	0,71	0,01	0,14	0,01	0,00	0,00	0,00	0,00	0,01	0,22
16	0,33	0,07	0,63	0,06	5,86	13,19	0,29	13,92	0,37	1,28	0,04	0,02	0,01	0,05	0,01	0,00	0,12
17	0,09	0,03	0,16	0,02	0,48	0,06	0,14	4,45	0,08	1,94	0,19	0,01	0,00	0,01	0,22	0,12	0,00
18	0,02	0,01	0,04	0,00	0,12	0,01	0,04	1,21	0,02	0,32	0,03	0,00	0,00	0,00	0,20	0,03	1,30
19	0,26	0,07	0,43	0,04	1,39	0,12	0,44	19,72	0,11	1,45	0,09	0,01	0,01	0,04	0,28	0,25	0,75
20	0,13	0,03	0,23	0,02	1,02	0,13	0,18	8,33	0,21	5,25	0,09	0,01	0,00	0,02	0,10	0,26	1,06
21	0,42	0,19	0,70	0,08	0,61	0,05	2,57	9,34	0,03	0,32	0,03	0,01	0,01	0,06	0,01	0,11	0,09
22	1,87	0,20	2,31	0,22	14,53	0,47	0,88	93,55	0,12	0,78	0,03	0,04	0,02	0,12	0,02	0,89	0,13
23	0,16	0,06	0,32	0,03	0,74	0,10	0,29	6,46	0,12	2,70	14,17	0,01	0,01	0,03	0,08	0,22	1,56
24	0,28	0,09	0,53	0,05	1,26	0,16	0,49	11,33	0,20	4,37	6,28	0,02	0,01	0,05	0,18	0,36	3,48
25	0,81	0,13	2,76	0,17	0,79	0,08	0,32	5,07	0,02	0,17	0,01	0,05	0,03	0,16	0,00	0,18	0,03
26	7,92	0,70	31,77	1,58	2,99	0,26	1,74	23,22	0,09	0,64	0,04	0,15	0,10	0,78	0,01	0,58	0,11
27	0,19	0,10	0,69	0,05	0,23	0,03	0,20	1,71	0,01	0,10	0,01	2,31	2,77	0,17	0,00	0,07	0,02
28	0,06	0,04	0,23	0,02	0,07	0,01	0,07	0,54	0,00	0,03	0,00	0,23	0,25	0,07	0,00	0,02	0,01
29	0,36	0,26	1,41	0,11	0,37	0,04	0,43	2,96	0,02	0,16	0,01	0,60	0,56	0,50	0,00	0,10	0,03
30	0,62	0,41	2,88	0,21	0,54	0,06	0,62	4,33	0,03	0,21	0,01	0,72	0,56	1,41	0,00	0,14	0,04
31	3,93	1,65	18,47	1,22	3,60	0,40	3,00	26,93	0,16	1,25	0,08	3,85	2,32	4,34	0,03	0,96	0,22
32	0,55	0,25	3,53	0,21	0,35	0,04	0,38	1,83	0,01	0,11	0,01	0,13	0,09	1,33	0,00	0,08	0,02
33	0,24	0,04	0,73	0,05	0,34	0,04	0,11	2,99	0,01	0,08	0,00	0,02	0,01	0,05	0,00	0,09	0,01
34	0,61	0,09	1,71	0,12	0,86	0,09	0,26	4,96	0,03	0,17	0,01	0,04	0,02	0,10	0,00	0,21	0,03

Tabela B.4: Matriz de tráfego usada para o cenário de tráfego não-uniforme para a rede ARNES (conclusão).

1	0,02	0,26	0,13	0,42	1,87	0,16	0,28	0,81	7,92	0,19	0,06	0,36	0,62	3,93	0,55	0,24	0,61
2	0,01	0,07	0,03	0,19	0,20	0,06	0,09	0,13	0,70	0,10	0,04	0,26	0,41	1,65	0,25	0,04	0,09
3	0,04	0,43	0,23	0,70	2,31	0,32	0,53	2,76	31,77	0,69	0,23	1,41	2,88	18,47	3,53	0,73	1,71
4	0,00	0,04	0,02	0,08	0,22	0,03	0,05	0,17	1,58	0,05	0,02	0,11	0,21	1,22	0,21	0,05	0,12
5	0,12	1,39	1,02	0,61	14,53	0,74	1,26	0,79	2,99	0,23	0,07	0,37	0,54	3,60	0,35	0,34	0,86
6	0,01	0,12	0,13	0,05	0,47	0,10	0,16	0,08	0,26	0,04	0,01	0,04	0,06	0,40	0,04	0,09	0,09
7	0,04	0,44	0,18	2,57	0,88	0,29	0,49	0,32	1,74	0,20	0,07	0,43	0,62	3,00	0,38	0,11	0,26
8	1,21	19,72	8,33	9,34	93,55	6,46	11,33	5,07	23,22	1,71	0,54	2,96	4,33	26,93	2,83	1,99	4,96
9	0,02	0,11	0,21	0,03	0,12	0,12	0,20	0,02	0,09	0,01	0,00	0,02	0,03	0,16	0,01	0,01	0,03
10	0,32	1,45	5,25	0,32	0,78	2,70	4,37	0,17	0,64	0,10	0,03	0,16	0,21	1,25	0,11	0,08	0,17
11	0,03	0,09	0,09	0,03	0,03	14,17	6,28	0,01	0,04	0,01	0,00	0,01	0,01	0,08	0,01	0,00	0,01
12	0,00	0,01	0,01	0,01	0,04	0,01	0,02	0,05	0,15	2,31	0,23	0,60	0,72	3,85	0,13	0,02	0,04
13	0,00	0,01	0,00	0,01	0,02	0,01	0,01	0,03	0,10	2,77	0,25	0,56	0,56	2,32	0,09	0,01	0,02
14	0,00	0,04	0,02	0,06	0,12	0,03	0,05	0,16	0,78	0,17	0,07	0,50	1,41	4,34	1,33	0,05	0,10
15	0,20	0,28	0,10	0,01	0,02	0,08	0,18	0,00	0,01	0,00	0,00	0,00	0,00	0,03	0,00	0,00	0,00
16	0,03	0,25	0,26	0,11	0,89	0,22	0,36	0,18	0,58	0,07	0,02	0,10	0,14	0,96	0,08	0,09	0,21
17	1,30	0,75	1,06	0,09	0,13	1,56	3,48	0,03	0,11	0,02	0,01	0,03	0,04	0,22	0,02	0,01	0,03
18	0,00	0,29	0,19	0,03	0,03	0,28	0,64	0,01	0,03	0,00	0,00	0,01	0,01	0,05	0,01	0,00	0,01
19	0,29	0,00	0,97	0,32	0,41	0,64	1,24	0,07	0,29	0,04	0,01	0,06	0,09	0,50	0,05	0,03	0,07
20	0,19	0,97	0,00	0,10	0,23	0,64	1,16	0,04	0,17	0,02	0,01	0,04	0,05	0,31	0,03	0,02	0,04
21	0,03	0,32	0,10	0,00	0,29	0,17	0,31	0,08	0,37	0,05	0,02	0,10	0,13	0,66	0,07	0,03	0,07
22	0,03	0,41	0,23	0,29	0,00	0,22	0,37	0,50	2,30	0,11	0,03	0,18	0,28	1,91	0,20	0,20	0,53
23	0,28	0,64	0,64	0,17	0,22	0,00	154,92	0,06	0,22	0,04	0,01	0,07	0,08	0,47	0,04	0,02	0,05
24	0,64	1,24	1,16	0,31	0,37	154,92	0,00	0,10	0,36	0,06	0,02	0,11	0,14	0,78	0,07	0,04	0,09
25	0,01	0,07	0,04	0,08	0,50	0,06	0,10	0,00	7,49	0,14	0,04	0,21	0,41	4,26	0,40	2,92	6,83
26	0,03	0,29	0,17	0,37	2,30	0,22	0,36	7,49	0,00	0,43	0,14	0,76	1,49	12,55	1,68	1,36	3,68
27	0,00	0,04	0,02	0,05	0,11	0,04	0,06	0,14	0,43	0,00	1,79	2,88	1,90	7,27	0,34	0,06	0,11
28	0,00	0,01	0,01	0,02	0,03	0,01	0,02	0,04	0,14	1,79	0,00	4,18	0,93	2,07	0,12	0,02	0,03
29	0,01	0,06	0,04	0,10	0,18	0,07	0,11	0,21	0,76	2,88	4,18	0,00	7,27	10,37	0,73	0,08	0,16
30	0,01	0,09	0,05	0,13	0,28	0,08	0,14	0,41	1,49	1,90	0,93	7,27	0,00	37,23	2,93	0,15	0,28
31	0,05	0,50	0,31	0,66	1,91	0,47	0,78	4,26	12,55	7,27	2,07	10,37	37,23	0,00	22,86	1,54	2,75
32	0,01	0,05	0,03	0,07	0,20	0,04	0,07	0,40	1,68	0,34	0,12	0,73	2,93	22,86	0,00	0,12	0,24
33	0,00	0,03	0,02	0,03	0,20	0,02	0,04	2,92	1,36	0,06	0,02	0,08	0,15	1,54	0,12	0,00	6,52
34	0,01	0,07	0,04	0,07	0,53	0,05	0,09	6,83	3,68	0,11	0,03	0,16	0,28	2,75	0,24	6,52	0,00

Tabela B.5: Matriz de tráfego usada para o cenário de tráfego não-uniforme para a rede FTLG (continua).

1	0,00	0,13	0,23	0,18	0,31	0,06	0,34	0,28	0,40	0,54	0,10	0,13	1,61	0,15	0,43	0,02	0,42	0,17	0,21
2	9,97	0,00	0,98	1,25	1,61	0,48	2,81	2,82	5,70	9,43	0,51	0,58	4,53	0,33	1,56	0,05	0,38	0,49	0,38
3	0,13	0,98	0,00	1,71	0,33	0,10	0,14	0,24	0,18	0,14	0,14	0,14	0,36	0,04	0,23	0,00	0,06	0,02	0,03
4	0,23	2,09	1,71	0,00	1,21	0,51	0,37	0,36	0,55	0,38	0,16	0,15	0,46	0,04	0,26	0,01	0,08	0,03	0,05
5	0,18	1,25	0,52	1,21	0,00	1,43	0,75	0,40	0,75	0,36	0,26	0,27	0,39	0,04	0,26	0,01	0,09	0,03	0,04
6	0,31	1,61	0,33	0,51	1,43	0,00	0,23	0,51	0,54	0,30	1,21	0,71	0,80	0,09	0,63	0,02	0,22	0,08	0,10
7	0,06	0,48	0,10	0,37	0,75	0,23	0,00	0,23	0,82	0,22	0,13	0,04	0,10	0,01	0,06	0,00	0,02	0,01	0,01
8	0,34	2,81	0,14	0,36	0,40	0,51	0,23	0,00	2,45	1,99	1,44	0,10	0,09	0,36	0,03	0,18	0,10	0,04	0,05
9	0,28	2,82	0,24	0,83	0,75	0,54	0,82	2,45	0,00	3,16	1,23	0,11	0,37	0,03	0,19	0,01	0,09	0,03	0,04
10	0,40	5,70	0,18	0,55	0,36	0,35	0,22	1,99	3,16	0,00	7,64	0,09	0,39	0,03	0,18	0,01	0,09	0,04	0,05
11	0,54	9,43	0,14	0,38	0,26	0,30	0,13	1,44	1,23	7,64	0,00	0,08	0,41	0,03	0,18	0,01	0,10	0,04	0,05
12	0,10	0,51	0,14	0,16	0,27	1,21	0,04	0,10	0,11	0,09	0,08	0,00	2,47	0,43	0,06	0,50	0,01	0,09	0,04
13	0,13	0,58	0,14	0,15	0,21	0,71	0,04	0,09	0,11	0,09	0,08	2,47	0,00	0,67	0,11	1,08	0,02	0,13	0,05
14	1,61	4,53	0,36	0,46	0,39	0,80	0,10	0,36	0,37	0,39	0,41	0,43	0,67	0,00	3,38	6,28	1,74	1,11	2,37
15	0,10	0,33	0,04	0,04	0,04	0,09	0,01	0,03	0,03	0,03	0,06	0,11	3,38	0,00	3,24	0,12	0,15	0,08	0,15
16	0,43	1,56	0,23	0,26	0,26	0,63	0,06	0,18	0,19	0,18	0,50	1,08	6,28	3,24	0,00	0,26	0,59	0,29	0,47
17	0,02	0,05	0,00	0,01	0,01	0,02	0,00	0,01	0,01	0,01	0,01	0,01	0,36	0,12	0,26	0,00	0,04	0,02	0,05
18	0,42	0,93	0,06	0,08	0,09	0,22	0,10	0,09	0,09	0,10	0,09	0,13	1,74	0,15	0,59	0,04	0,00	3,03	1,22
19	0,17	0,38	0,02	0,03	0,03	0,08	0,01	0,04	0,04	0,04	0,04	0,05	1,11	0,08	0,29	0,02	3,03	0,00	2,37
20	0,21	0,49	0,03	0,05	0,04	0,10	0,01	0,05	0,04	0,05	0,05	0,07	2,37	0,15	0,47	0,05	1,22	2,37	0,00
21	0,03	0,07	0,01	0,01	0,01	0,02	0,00	0,01	0,01	0,01	0,01	0,01	0,11	0,01	0,05	0,00	0,40	0,06	0,05
22	0,01	0,03	0,00	0,00	0,00	0,02	0,00	0,01	0,00	0,00	0,01	0,01	0,05	0,01	0,03	0,00	0,46	0,01	0,01
23	0,01	0,04	0,00	0,01	0,01	0,03	0,00	0,01	0,01	0,01	0,01	0,01	0,04	0,00	0,03	0,00	0,03	0,01	0,01
24	0,04	0,15	0,02	0,03	0,04	0,18	0,01	0,03	0,02	0,02	0,08	0,09	0,15	0,02	0,14	0,00	0,07	0,02	0,03
25	0,03	0,08	0,01	0,01	0,01	0,04	0,00	0,01	0,01	0,01	0,01	0,02	0,08	0,01	0,05	0,00	0,09	0,02	0,02
26	0,30	0,71	0,04	0,07	0,07	0,20	0,02	0,09	0,07	0,08	0,08	0,10	0,87	0,08	0,36	0,02	4,05	0,49	0,35
27	0,03	0,07	0,00	0,01	0,01	0,02	0,00	0,01	0,01	0,01	0,01	0,01	0,06	0,01	0,02	0,00	0,13	0,02	0,02
28	0,01	0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,06	0,01	0,01
29	2,79	5,45	0,30	0,44	0,41	0,99	1,11	0,51	0,45	0,49	0,54	0,42	13,40	0,85	3,01	0,21	23,52	26,46	13,80
30	0,38	0,63	0,03	0,04	0,04	0,09	0,01	0,05	0,04	0,05	0,06	0,04	1,24	0,07	0,24	0,01	0,80	0,49	0,52
31	0,25	0,41	0,02	0,03	0,03	0,06	0,01	0,03	0,03	0,03	0,04	0,02	0,75	0,04	0,15	0,01	0,62	0,35	0,33
32	0,20	0,34	0,02	0,02	0,02	0,05	0,01	0,03	0,03	0,03	0,02	0,03	0,55	0,03	0,13	0,01	0,71	0,32	0,26
33	0,43	0,69	0,03	0,05	0,04	0,11	0,01	0,07	0,05	0,06	0,07	0,04	0,62	0,04	0,18	0,01	0,86	0,24	0,20
34	0,06	0,10	0,00	0,01	0,01	0,01	0,00	0,01	0,01	0,01	0,01	0,01	0,11	0,01	0,03	0,00	0,13	0,05	0,04
35	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,01	0,00	0,00
36	0,02	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,01	0,00	0,00
37	0,21	0,33	0,01	0,02	0,02	0,05	0,01	0,03	0,03	0,04	0,02	0,02	0,22	0,02	0,07	0,00	0,22	0,06	0,06
38	0,22	0,22	0,01	0,01	0,01	0,02	0,00	0,02	0,01	0,02	0,01	0,01	0,06	0,00	0,02	0,00	0,03	0,01	0,01
39	0,07	0,13	0,00	0,01	0,01	0,02	0,00	0,02	0,01	0,01	0,01	0,02	0,04	0,00	0,02	0,00	0,03	0,01	0,01
40	51,68	1,19	0,02	0,03	0,02	0,04	0,01	0,04	0,05	0,07	0,01	0,01	0,19	0,01	0,05	0,00	0,05	0,03	0,03
41	1,57	200,00	0,10	0,21	0,13	0,18	0,05	0,30	0,28	0,55	0,91	0,06	0,54	0,04	0,18	0,01	0,12	0,05	0,06

Tabela B.6: Matriz de tráfego usada para o cenário de tráfego não-uniforme para a rede FTLG (conclusão).

1	0,03	0,01	0,01	0,04	0,03	0,30	0,03	0,29	0,38	0,25	0,20	0,43	0,06	0,01	0,02	0,21	0,22	0,07	51,68	41
2	0,07	0,03	0,04	0,15	0,08	0,71	0,07	5,45	0,63	0,41	0,34	0,69	0,10	0,01	0,03	0,33	0,22	0,13	1,19	1,57
3	0,01	0,00	0,00	0,02	0,01	0,04	0,00	0,30	0,03	0,02	0,02	0,03	0,00	0,00	0,00	0,01	0,01	0,00	0,02	200,00
4	0,01	0,00	0,01	0,03	0,01	0,07	0,01	0,44	0,04	0,03	0,02	0,05	0,01	0,00	0,00	0,02	0,01	0,01	0,03	0,21
5	0,01	0,00	0,01	0,04	0,01	0,07	0,01	0,41	0,04	0,03	0,02	0,04	0,01	0,00	0,00	0,00	0,01	0,01	0,02	0,13
6	0,02	0,02	0,03	0,18	0,04	0,20	0,02	0,99	0,09	0,06	0,05	0,11	0,01	0,00	0,00	0,05	0,02	0,02	0,04	0,18
7	0,00	0,00	0,00	0,01	0,00	0,02	0,00	0,11	0,01	0,01	0,01	0,01	0,01	0,00	0,00	0,01	0,00	0,00	0,01	0,05
8	0,01	0,01	0,01	0,03	0,01	0,09	0,01	0,51	0,05	0,03	0,03	0,07	0,01	0,00	0,00	0,03	0,02	0,02	0,04	0,30
9	0,01	0,00	0,01	0,03	0,01	0,07	0,01	0,45	0,04	0,03	0,03	0,05	0,01	0,00	0,00	0,03	0,01	0,01	0,04	0,28
10	0,01	0,00	0,01	0,02	0,01	0,08	0,01	0,49	0,05	0,03	0,03	0,06	0,01	0,00	0,00	0,03	0,02	0,01	0,05	0,55
11	0,01	0,01	0,01	0,02	0,01	0,08	0,01	0,54	0,06	0,04	0,03	0,07	0,01	0,00	0,00	0,04	0,02	0,02	0,07	0,91
12	0,01	0,01	0,01	0,08	0,02	0,08	0,01	0,42	0,04	0,02	0,02	0,04	0,01	0,00	0,00	0,02	0,01	0,00	0,01	0,06
13	0,01	0,01	0,01	0,09	0,02	0,10	0,01	0,59	0,05	0,03	0,03	0,05	0,01	0,00	0,00	0,02	0,01	0,01	0,02	0,06
14	0,11	0,05	0,04	0,15	0,08	0,87	0,06	13,40	1,24	0,75	0,55	0,62	0,11	0,01	0,01	0,22	0,06	0,04	0,19	0,54
15	0,01	0,01	0,00	0,02	0,01	0,08	0,01	0,85	0,07	0,04	0,03	0,04	0,01	0,00	0,00	0,02	0,00	0,00	0,01	0,04
16	0,05	0,03	0,03	0,14	0,05	0,36	0,02	3,01	0,24	0,15	0,13	0,18	0,03	0,00	0,00	0,07	0,02	0,02	0,05	0,18
17	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,21	0,01	0,01	0,01	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,01
18	0,40	0,06	0,03	0,07	0,09	4,05	0,13	23,52	0,80	0,62	0,71	0,86	0,13	0,01	0,01	0,22	0,03	0,03	0,05	0,12
19	0,06	0,01	0,01	0,02	0,02	0,49	0,02	26,46	0,49	0,35	0,32	0,24	0,05	0,00	0,00	0,06	0,01	0,01	0,02	0,05
20	0,05	0,01	0,01	0,02	0,02	0,35	0,02	13,80	0,52	0,33	0,26	0,20	0,04	0,00	0,00	0,06	0,01	0,01	0,03	0,06
21	0,00	0,03	0,01	0,01	0,02	0,66	0,01	0,53	0,03	0,02	0,02	0,04	0,01	0,00	0,00	0,01	0,00	0,00	0,00	0,01
22	0,03	0,00	0,01	0,01	0,03	0,08	0,00	0,14	0,01	0,01	0,01	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
23	0,01	0,01	0,00	0,04	0,04	0,03	0,00	0,09	0,01	0,00	0,00	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
24	0,01	0,01	0,04	0,00	0,03	0,07	0,00	0,24	0,02	0,01	0,01	0,02	0,00	0,00	0,00	0,01	0,00	0,00	0,01	0,02
25	0,02	0,03	0,04	0,03	0,00	0,15	0,01	0,24	0,02	0,01	0,01	0,01	0,03	0,00	0,00	0,01	0,00	0,00	0,00	0,01
26	0,66	0,08	0,03	0,07	0,15	0,00	0,29	5,37	0,31	0,24	0,27	0,65	0,07	0,01	0,01	0,21	0,02	0,03	0,04	0,09
27	0,01	0,00	0,00	0,00	0,01	0,29	0,00	0,34	0,03	0,02	0,02	0,11	0,01	0,00	0,00	0,05	0,00	0,00	0,00	0,01
28	0,00	0,00	0,00	0,00	0,00	0,11	0,04	0,14	0,01	0,01	0,01	0,05	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00
29	0,53	0,14	0,09	0,24	0,24	5,37	0,34	19,39	15,05	13,56	13,56	4,73	1,26	0,12	0,06	1,10	0,17	0,13	0,35	0,68
30	0,03	0,01	0,01	0,02	0,02	0,31	0,03	0,01	19,39	21,63	2,42	0,53	0,21	0,02	0,01	0,12	0,02	0,01	0,05	0,08
31	0,02	0,01	0,00	0,01	0,01	0,24	0,02	15,05	21,63	2,42	3,91	0,45	0,20	0,02	0,01	0,10	0,01	0,01	0,03	0,05
32	0,02	0,01	0,00	0,01	0,01	0,27	0,02	13,56	2,42	3,91	0,00	0,61	0,37	0,02	0,01	0,11	0,01	0,01	0,03	0,04
33	0,04	0,01	0,01	0,02	0,03	0,65	0,11	4,73	0,53	0,45	0,61	0,00	0,48	0,08	0,04	1,43	0,05	0,04	0,06	0,09
34	0,01	0,00	0,00	0,00	0,00	0,07	0,01	1,26	0,21	0,20	0,37	0,48	0,00	0,05	0,00	0,06	0,01	0,01	0,01	0,01
35	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,12	0,02	0,02	0,02	0,08	0,05	0,00	0,00	0,01	0,00	0,00	0,00	0,00
36	0,00	0,00	0,00	0,00	0,00	0,01	0,00	0,06	0,01	0,01	0,01	0,04	0,00	0,00	0,00	0,02	0,00	0,00	0,00	0,00
37	0,01	0,00	0,00	0,01	0,01	0,21	0,05	1,10	1,12	1,10	1,11	1,43	0,06	0,01	0,02	0,00	0,03	0,04	0,03	0,04
38	0,00	0,00	0,00	0,00	0,00	0,02	0,00	0,17	0,02	0,01	0,01	0,05	0,01	0,00	0,00	0,03	0,00	0,02	0,03	0,03
39	0,00	0,00	0,00	0,00	0,00	0,03	0,00	0,13	0,01	0,01	0,01	0,04	0,00	0,00	0,00	0,04	0,02	0,00	0,01	0,02
40	0,00	0,00	0,00	0,01	0,00	0,04	0,00	0,35	0,05	0,03	0,03	0,06	0,01	0,00	0,00	0,03	0,03	0,01	0,00	0,19
41	0,01	0,00	0,00	0,02	0,01	0,09	0,01	0,68	0,08	0,05	0,04	0,09	0,01	0,00	0,00	0,04	0,03	0,02	0,19	0,00