

Universidade Federal do Pará
Centro Tecnológico
Programa de Pós-Graduação em Engenharia Elétrica

Metodologia de Identificação de Topologias em Sistemas
DSL Baseada em Otimização Evolucionária
Multiobjetivo

CLAUDOMIRO DE SOUZA DE SALES JUNIOR

Orientador:

PROF. DR. JOÃO CRISÓSTOMO WEYL ALBUQUERQUE COSTA

Belém
2009.

Metodologia de Identificação de Topologias em Sistemas DSL Baseada em Otimização Evolucionária Multiobjetivo

CLAUDOMIRO DE SOUZA DE SALES JUNIOR

Orientador:

PROF. DR. JOÃO CRISÓSTOMO WEYL ALBUQUERQUE COSTA

Proposta de Tese de Doutorado submetida à Banca Examinadora do Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal do Pará como pré-requisito para obtenção do título de “Doutor em Engenharia Elétrica”.

Universidade Federal do Pará

Belém
2009

Metodologia de Identificação de Topologias em Sistemas DSL Baseada em Otimização Evolucionária Multiobjetivo

Claudio Miros de Souza de Sales Junior

Banca examinadora

.....
Prof. Dr. João Crisóstomo Weyl Albuquerque Costa (UFPA) - Orientador

.....
Prof. Dr. Gervásio Protásio dos Santos Cavalcante (UFPA) - Membro

.....
Prof. Dr. Aldebaro Barreto da Rocha Klautau Júnior (UFPA) - Membro

.....
Prof. Dr. Agostinho Luiz da Silva Castro (UFPA) - Membro

.....
Prof. Dr. Paulo Augusto Valente Ferreira (UNICAMP) - Membro

.....
Prof. Dr. Marcelo Eduardo Vieira Segatto (UFES) - Membro

.....
Prof. Dr. Nandamud Lankalapalli Vijaykumar (INPE) - Membro

.....
Prof. Dr. Solon Venâncio de Carvalho (INPE) - Membro

Visto:

.....
Prof. Dr. Evaldo Gonçalves Pelaes

Coordenador do PPGEE/CT/UFPA

Resumo

A análise de desempenho da tecnologia de linhas digitais do assinante (*Digital Subscriber Line* -DSL), que utiliza a infraestrutura cabeada da rede telefônica existente, é de vital importância para os provedores de serviço DSL. Esse tipo de análise exige a predição precisa da topologia do enlace com a identificação dos parâmetros do cabo, isto é, a determinação do comprimento e a bitola (diâmetro do fio) de todas as seções do enlace telefônico (incluindo as ramificações, chamadas de *bridged-taps*). A identificação da topologia do enlace permitirá aos operadores qualificar precisamente um enlace para o serviço DSL, para atualizar e reorganizar os dados da sua rede telefônica. Este trabalho apresenta um método para identificação da topologia de enlaces baseado nas novas funcionalidades de diagnóstico de enlace das recomendações ITU-T de modems ADSL2 e ADSL2+. A metodologia desenvolvida emprega medidas de uma porta (*single-ended line testing* - SELT) e medidas de duas portas (*double-ended line testing* - DELT) para identificar a topologia dos enlaces sob teste através de um algoritmo genético (AG) de otimização multiobjetivo, uma técnica local de otimização e métodos analíticos. As medidas de SELT e DELT são usadas para compor duas funções custo para os processos de otimização e também são usadas pelas técnicas analíticas para obter informações sobre o enlace. O procedimento de otimização consiste de um processo híbrido composto de uma versão especializada do AG por ordenamento por não-dominância (*Non-Dominated Sorting Genetic Algorithm*, NSGA-II) e a técnica local **baseada** no método de *Levenberg-Marquardt* para pesquisar a topologia candidata que melhor alcança o critério definido de otimização. Além disso, uma estratégia de controle adaptativo para melhorar a convergência do algoritmo genético é proposta. Os resultados obtidos mostram que o método desenvolvido é eficiente e preciso.

Abstract

Performance analysis of Digital Subscriber Line (DSL) technologies, which are implemented on existing telephone subscriber loops, is of vital importance to DSL service providers. This type of analysis requires accurate prediction of the loop topology with precise identification of the cable parameters, i.e., the determination of the length and the gauge of all loop sections (including bridged taps). Loop topology identification will allow operators to more accurately qualify a loop for DSL service and to update and reorder telephone-company loop records, which can be accessed to support engineering, provisioning, and maintenance operations. This work presents a method for loop topology identification based on the new loop diagnostic functionality of the ITU-T standard for ADSL2 and ADSL2+ modems. The proposed methodology employs single-ended line testing (SELT) and double-ended line testing (DELT) measurements to identify the topology of the loop under test through a multi-objective genetic algorithm, a local technique and analytical methods. The SELT and DELT measurements are used to compose two cost functions for the optimization processes and are also used by local techniques to obtain information about loop. The optimization procedure consists of a specialized version of the non-dominated sorting genetic algorithm-II (NSGA-II) and local technique based on *Levenberg-Marquardt* method that seeks the candidate topology that best matches the defined criterion of the optimization. Besides, an adaptive parameter control strategy to improve the GA convergence is proposed. Simulation results show that the developed method is efficient and accurate.

Sumário

Lista de Figuras	iv
Glossário	vii
1 Introdução	1
1.1 Linha Digital de Assinante	1
1.2 Contexto e Motivação	4
1.2.1 Métodos de Identificação de Linha	5
1.3 Objetivos do Trabalho	7
1.4 Contribuições da Tese	9
1.5 Estrutura do Trabalho	12
2 Análise e Modelagem do Par Trançado	13
2.1 Análise do par trançado	13
2.1.1 Características do par trançado	13
2.1.2 Teoria de Matriz ABCD	16
2.2 Modelos de linha	18
2.2.1 Modelo RLGC analítico	20
2.2.2 Modelo RLCG pela contribuição ANSI T1E1.4	20
2.2.3 Modelo RLCG pela contribuição ETSI	21
2.2.4 Modelo RLCG pela contribuição MAR	22
2.2.5 Modelo VUB	23
2.3 Conclusões do Capítulo	25
3 Algoritmo Genético Multiobjetivo	26
3.1 Algoritmos Genéticos	26
3.1.1 Conceito Básico	26
3.1.1.1 Tipos de Codificação para os Algoritmos Genéticos	27
3.1.2 Operadores Genéticos	28

3.1.2.1	Seleção	28
3.1.2.2	Cruzamento	31
3.1.2.3	Mutação	33
3.1.2.4	Elitismo	34
3.1.3	Parâmetros genéticos	34
3.2	Otimização Multiobjetivo	35
3.2.1	Problema de Otimização Multiobjetivo	36
3.2.2	Soluções Ótimas de Pareto	37
3.2.3	Dominância e Otimalidade de Pareto	38
3.2.4	Objetivos em Otimização Multiobjetivo	39
3.2.5	Diferença entre Otimização Mono-objetivo e Multiobjetivo	39
3.2.6	Métodos Clássicos de Otimização Multiobjetivo	40
3.3	Algoritmos Evolutivos para Otimização Multiobjetivo	42
3.3.1	<i>Vector Evaluated Genetic Algorithm</i> - VEGA	43
3.3.2	<i>Multi-objective Optimization Genetic Algorithm</i> - MOGA	43
3.3.3	<i>Niched-Pareto Genetic Algorithm</i> - NPGA	44
3.3.4	<i>Strength Pareto Evolutionary Algorithm</i> - SPEA	44
3.3.5	<i>Non-Dominated Sorting Genetic Algorithm</i>	45
3.3.5.1	<i>Non-Dominated Sorting Genetic Algorithm</i> - II	46
3.3.5.2	Distância de Multidão	47
3.3.5.3	Operador de comparação de multidão	48
3.4	Conclusões do Capítulo	50
4	Identificação de Topologias Usando Computação Evolucionária Multiobje-	
	tivo - TIMEC	51
4.1	Ferramenta de Qualificação Identificação de enlaces DSL - LQ-SUITE	51
4.2	O método TIMEC	52
4.2.1	Otimização Baseada em AG	54
4.2.2	Otimização local e processo de realimentação de seções	58
4.2.3	Processo de Codificação e Decodificação	62
4.2.3.1	Codificação/decodificação do Comprimento e Bitola	62
4.2.3.2	Decodificação do Número de Seções	65
4.2.3.3	Decodificação do Tipo das Seções	65
4.2.3.4	Um Exemplo de Decodificação	66
4.2.4	Controle Auto-adaptativo dos Parâmetros do AG	67
4.3	Ferramenta de Medição Automática	70

4.4	Conclusões do Capítulo	71
5	Técnicas Analíticas de Pré-processamento	73
5.1	Técnicas de Pré-processamento para Obtenção de Informação <i>a priori</i>	73
5.1.1	Extrator de Características baseado em Wavelet - WFE	74
5.1.2	IRA - <i>Impulse Response Analysis</i>	77
5.1.3	First Section Finder (FSF)	84
5.2	Conclusões do Capítulo	89
6	Resultados	90
6.1	Comparação de TIMEC e <i>SELT-tdr</i>	90
6.2	Resultados do LQ-SUITE	96
6.2.1	Dados Simulados	96
6.2.1.1	Resultados do IRA	96
6.2.1.2	Resultados do FSF	99
6.2.1.3	Resultados LQ-SUITE	101
6.2.2	Dados Medidos	101
6.2.2.1	Resultados do IRA	102
6.2.2.2	Resultados do FSF	103
6.2.2.3	Resultados do LQ-SUITE	105
6.3	Conclusões do Capítulo	108
	Conclusão	110
	Referências	114
	Publicações Relacionadas	120
A	Setup de Medição de Função de Transferência H e Parâmetro de Espalhamento S_{11}	121
A.1	Lista de Equipamentos Utilizados	121
A.2	Configuração dos Equipamentos de Medição	122
A.3	Breve Descrição do Processo de Medição	122
B	Enlaces ETSI	124

Lista de Figuras

1.1	Sistema DSL típico, com o modem do usuário à esquerda conectado à rede metálica e o <i>Central Office</i> à direita a partir do bastidor de <i>splitters</i> conectado à rede metálica.	3
1.2	Parâmetros do enlace Θ a serem determinados. Nesse exemplo, a terceira seção é uma <i>bridged-tap</i> enquanto todas as outras são seriais.	4
1.3	A metodologia DELTI aplicada em um sistema DSL.	9
1.4	Visão resumida das várias áreas de ação e possíveis contribuições associadas a esta proposta de tese.	11
2.1	Modelo de linha de transmissão por unidade de comprimento, mostrando a indutância L , a resistência R , a capacitância C e a condutância G	15
2.2	Representação de um enlace DSL. Nesse enlace, a terceira seção é uma <i>bridged-tap</i> enquanto as outras seções são seriais. O número total de seções é n_s	16
2.3	Quadripolo ABCD.	16
2.4	A seção de linha infinitesimal.	20
2.5	Seção transversal de uma linha de 2 fios paralelos.	23
3.1	Diagrama do algoritmo genético.	27
3.2	Representação do espaço de variáveis de decisão e o correspondente espaço de objetivos.	37
3.3	Conceito de soluções ótimas de Pareto.	38
3.4	Esquema do modelo NSGA-II [1].	47
3.5	Algoritmo para cálculo da distância de multidão.	48
3.6	Procedimento para cálculo da distância de multidão [1].	48
4.1	Parâmetros do enlace Θ a serem determinados. Nesse exemplo, a terceira seção é uma <i>bridged-tap</i> enquanto todas as outras são seriais.	52
4.2	LQ-SUITE e seus módulos: TIMEC, IRA e FSF.	53
4.3	Diagrama de bloco do TIMEC com apenas a otimização global.	55

4.4	Diagrama de bloco do TIMEC com otimização global e local.	59
4.5	Visão expandida do bloco \mathcal{H} de TIMEC.	60
4.6	a) Enlace obtido pelo AG na n -ésima geração e que serve de estimativa inicial a técnica local de otimização b) Enlace após a k -ésima execução do bloco \mathcal{H} , onde $k = n/N_{\mathcal{H}}$. Apenas o comprimento de cada seção é otimizado pela técnica de otimização local.	61
4.7	Processo de codificação (\mathcal{G}) e decodificação (\mathcal{G}^{-1}) para um enlace com o número máximo de seções N_s igual a 4, cada uma com comprimento l^i e bitola g^i	63
4.8	Exemplo de um cromossomo para ilustrar o processo de decodificação. Nesse caso o número máximo de seções N_s é igual a 4.	67
4.9	Diagrama de bloco da AMT.	71
5.1	Saída do WFE para três valores de k_s : 1, 1.5 e 2.	76
5.2	Resposta impulsiva com as bordas de seus picos identificados em vermelho.	78
5.3	Os picos possivelmente relacionados a <i>bridged-taps</i>	79
5.4	Determinação das bitolas das seções do enlace.	80
5.5	Picos possivelmente relacionados a <i>bridged-taps</i>	82
5.6	Comparação da curva medida e a obtida do enlace com as bitolas [0.4 0.5 0.4]mm.	83
5.7	Comparação da curva medida e a obtida do enlace com a <i>bridged-tap</i> adicional.	84
5.8	Método para determinar a bitola da primeira seção de um enlace.	86
5.9	Resposta impulsiva para o enlace de 500m e bitola 0.5mm.	87
5.10	Resposta impulsiva com três pontos identificados.	88
5.11	Resposta impulsiva com 4 pontos identificados.	89
6.1	Magnitude dos coeficientes de reflexão como função da frequência relacionados a descasamentos de impedância a mudanças de bitolas encontradas nos enlaces ETSI.	93
A.1	Cables used during the measurement process.	122
A.2	<i>Setup</i> de medição para função de transferência H e parâmetro de espalhamento S_{11} , sendo que para esse último apenas um balun é utilizado deixando a porta de saída em aberto.	123
B.1	ETSI#1.	124
B.2	ETSI#2.	124
B.3	ETSI#3.	124
B.4	ETSI#4.	125
B.5	ETSI#5.	125

B.6	ETSI#6.	125
B.7	ETSI#7.	125
B.8	ETSI#8.	125

Glossário

3GPP	-	<i>Third Generation Partnership Project</i>
8-PSK	-	<i>8-state Phase Shift Keying</i>
AIV	-	Algoritmo de Iteração de Valores
AJ	-	Adaptação Justa
AMPS	-	<i>Advanced Mobile Phone System</i>
ANSI	-	<i>American National Standards Institute</i>
B3G	-	<i>Beyond 3G</i>
BB	-	<i>Bandwidth Broker</i>
BLER	-	<i>Block Error Ratio</i>
RF	-	<i>Radio Frequency</i>
BMAP	-	<i>Batch Markovian Arrival Process</i>
BSC	-	<i>Base Station Controller</i>
CA	-	Com adaptação
CAC	-	Controle de Admissão de chamadas
CDMA	-	<i>Code Division Multiple Access</i>
CS	-	<i>Coding Scheme</i>
D-AMPS	-	<i>Digital AMPS</i>
DBMAP	-	<i>Discrete Batch Markovian Arrival Process</i>
DCA	-	<i>Dynamic Channel Assignment</i>
DMAP	-	<i>Discrete Markovian Arrival Process</i>
DSL	-	<i>Digital Subscriber Line</i>
DTBR	-	<i>Dual threshold bandwidth reservation</i>

ECSD	-	<i>Enhanced Circuit Switched Data</i>
EDGE	-	<i>Enhanced Data rates for GSM Evolution</i>
EGPRS	-	<i>Enhanced GPRS</i>
EM	-	<i>Estação Móvel</i>
ERB	-	<i>Estação Rádio Base</i>
ETSI	-	<i>European Telecommunications Standards Institute</i>
FCA	-	<i>Fixed Channel Assignment</i>
FCC	-	<i>Federal Communications Commissions</i>
FDD	-	<i>Frequency Division Duplex</i>
FDMA	-	<i>Frequency Division Multiple Access</i>
FTP	-	<i>File Transfer Protocol</i>
GMSK	-	<i>Gaussian Minimum Shift Keying</i>
GPRS	-	<i>General Packet Radio Service</i>
GPS	-	<i>General Position System</i>
GSM	-	<i>Global System for Mobile Communications</i>
HSCSD	-	<i>High Speed Circuit Switched Data</i>
IMT2000	-	<i>International Mobile Telecommunications 2000</i>
IP	-	<i>Internet Protocol</i>
IPP	-	<i>Interrupted Poisson Process</i>
IR	-	<i>Incremental Redundancy</i>
IS	-	<i>Interim Standard</i>
ITU	-	<i>International Telecommunications Union</i>
JDC	-	<i>Japanese Digital Cellular</i>
LA	-	<i>Link Adaptation</i>
LC	-	<i>Load Control</i>
LOTOS	-	<i>Language of Temporal Ordering Specification</i>
LQC	-	<i>Link Quality Control</i>
MAN	-	<i>Metropolitan Area Network</i>
MAP	-	<i>Markovian Arrival Process</i>
MMPP	-	<i>Markov Modulated Poisson Process</i>
MPEG	-	<i>Moving Pictures Experts Group</i>

NMT	-	<i>Nordic Mobile Telephone</i>
nG	-	Enésima Geração de telefonia móvel celular
NS	-	<i>Network Simulator</i>
OSI/ISO	-	<i>Open System Interconnection/ International Standards Organisation</i>
PDC	-	<i>Personal Digital Cellular</i>
PDCH	-	<i>Packet Data Channel</i>
PLC	-	<i>Power Line Communications</i>
PMD	-	Processo Markoviano de Decisão
PSMD	-	Processo Semi-Markoviano de Decisão
QoS	-	<i>Quality of Service</i>
RRM	-	<i>Radio Resource Management</i>
SA	-	Sem adaptação
SDL	-	<i>Specification and Description Language</i>
SDU	-	<i>Service Data Unit</i>
SIR	-	<i>Signal-to-Interference Ratio</i>
SLA	-	<i>Service Level Agreement</i>
SMS	-	<i>Short Message Service</i>
TCP	-	<i>Transmission Control Protocol</i>
TDD	-	<i>Time Division Duplex</i>
TDMA	-	<i>Time Division Multiple Access</i>
UMTS	-	<i>Universal Mobile Telecommunications System</i>
UTRA	-	<i>UMTS Terrestrial Radio Access</i>
VoIP	-	Voz sobre IP
WCDMA	-	<i>Wide band CDMA</i>
WLAN	-	<i>Wireless Local Area Network</i>
WWW	-	<i>World Wide Web</i>

Capítulo 1

Introdução

1.1 Linha Digital de Assinante

A planta telefônica, constituída de par trançado, está disponibilizada por toda região urbana, possuindo extensa capilaridade, abundância e abrangência, que a torna muito importante para prover métodos de acesso aos serviços de comunicações de dados, tais como: Internet, TV digital (HDTV), vídeo sob demanda (VoD), teleconferência, voz sobre IP (VoIP), sistema de segurança/monitoramento, entre outras.

A tecnologia de linha digital do assinante (em inglês *Digital Subscriber Line*, ou simplesmente DSL) torna possível a transmissão de altas de taxas de dados sobre a linha telefônica convencional. Isso é possível graças a utilização de técnicas sofisticadas de processamento de sinais, para mitigar efeitos causados pela atenuação, *crosstalk*, ruído, entre outros [2]. A principal vantagem dessa tecnologia sobre as outras tecnologias competidoras é a possibilidade de utilizar uma infra-estrutura pré-existente para transmissão de dados, tornando-a economicamente superior a qualquer outra tecnologia que necessite implantar uma nova infra-estrutura. Mesmo a transmissão via rádio necessita de uma infra-estrutura mínima. A partir da infra-estrutura já instalada é possível fazer facilmente testes de viabilidade para o serviço DSL, ao contrário da tecnologia de transmissão sem fio que pode envolver custos elevados para seu projeto e implantação [3]. Além disso, no sistema DSL não há compartilhamento de largura de banda entre os usuários do sistema como ocorre com a tecnologia que utiliza cabo coaxial (*cable modem*), na qual quanto maior o número de usuários utilizando o serviço, menor é a largura de banda disponível para cada usuário e assim menor a taxa de transmissão de cada usuário.

Vários tipos de tecnologias DSL têm surgido nos últimos anos dependendo da aplicação

e da padronização, podendo ser **classificadas** da seguinte forma [4, 5]:

- ADSL (*Asymmetric Digital Subscriber Lines*) - Essa tecnologia foi desenvolvida em 1987 pela *Bellcore Labs de Morristown*, Nova Jersey, como uma versão melhor que ISDN (*Integrated Services Digital Network*), para ser utilizada em “televisão interativa”. A comunicação é assimétrica e coexiste com o serviço tradicional de telefonia local, pelo mesmo par trançado. A taxa de transmissão pode alcançar até 8 Mbps, no canal *downstream* (da central para o assinante) e 640 Kbps no canal *upstream* (do assinante para a central);
- HDSL (*High Bit Rate Digital Subscriber Lines*) - Possui taxa de transmissão simétrica de 768 Kbps, 1,544 Mbps ou 2,048 Mbps. Essa tecnologia foi desenvolvida como uma variante rápida de ISDN. Dentre as aplicações, encontra-se a comunicação de canal E1 e T1, utilizando-se de modulação 2B1Q. Esse serviço não coexiste com o serviço de telefonia no mesmo par trançado e pode requerer até três pares trançados;
- ISDL é equivalente ao ISDN DL. São serviços simétricos, que utilizam a modulação 2B1Q, com taxas de transmissão de 56 Kbps, 64 Kbps, 128 Kbps e 144 Kbps;
- RDSL e RADSL têm como característica possuir canal reverso ao da tecnologia ADSL, onde a taxa *upstream* é maior que a taxa *downstream*. Esse tipo tem aplicação em provedor de Web;
- SDSL (*Symmetric Digital Subscriber Lines*) - Serviço simétrico com taxas entre 160 kbps e 2,048 Mbps, que é muito utilizado em operações comerciais. SDSL, diferentemente do ADSL, não pode coexistir com serviços de voz convencional;
- VDSL (*Very High Speed Digital Subscriber Line*) - As taxas de transmissões vão até 56 Mbps. As taxas de *downstream* são submúltiplas da velocidade das redes SONET/SDH (155,52 Mbps) e é indicada para aplicações em conjunto com acessos primários ópticos, como topologia FTTC (*Fiber to the Curb*) ou FFTN (*Fiber to the Neighborhood*).

A ADSL2 foi desenvolvida para alcançar melhor desempenho nas linhas de comprimentos longos na presença de interferência banda-básica [6]. Podem alcançar até 12 Mbps em canal *downstream* e 1 Mbps em canal *upstream*. Para a padronização ADSL2, especifica-se uma banda de frequência *downstream* de 1,1 MHz e 550 kHz para *upstream*. ADSL2+ alcançou reconhecimento pela ITU em janeiro de 2003, unindo ao padrão ADSL2 como família G.992.5 [7]. Em ADSL2+, especifica-se uma banda de frequência para *downstream* de até 2,2 MHz. Como consequência, resulta em uma significativa elevação na taxa de transmissão para *downstream*. A taxa *upstream* pode atingir até 1 Mbps, dependendo das condições da linha de assinante [5, 8].

O diagrama fim-a-fim de um sistema ADSL típico é mostrado na Figura 1.1. No lado do assinante deve haver um microfiltro ou *splitter* para separar as informações de voz e dados e enviá-los aos seus respectivos dispositivos na direção de *downstream*, e combinar as duas classes de informação na direção de *upstream*. O modem DSL pode consistir em um modem externo, um cartão de interface de rede, ou um roteador no caso de um pequeno escritório que necessite usar vários computadores. A conexão entre o equipamento de transmissão do assinante e a central (em inglês *Central Office* - CO) consiste em um cabo de par trançado que geralmente tem um comprimento de no máximo 5500 metros em aplicações ADSL [3]. No CO, um *splitter* é novamente necessário para combinar ou dividir as informações de voz e as informações de dados. A informação de voz é roteada a uma central telefônica (ou CPA - Central de Programa Armazenado) através de linhas de telefonia analógica tradicional (POTS - *Plain Old Telephone Service*) para a rede telefônica pública comutada (RTPC) e o dados são enviados ao DSLAM (do inglês, Digital Subscriber Line Access Multiplexer). O DSLAM agrega o tráfego da linha do assinante em um enlace de alta velocidade que está conectado ao equipamento de chaveamento do *backbone* e assim à Internet.

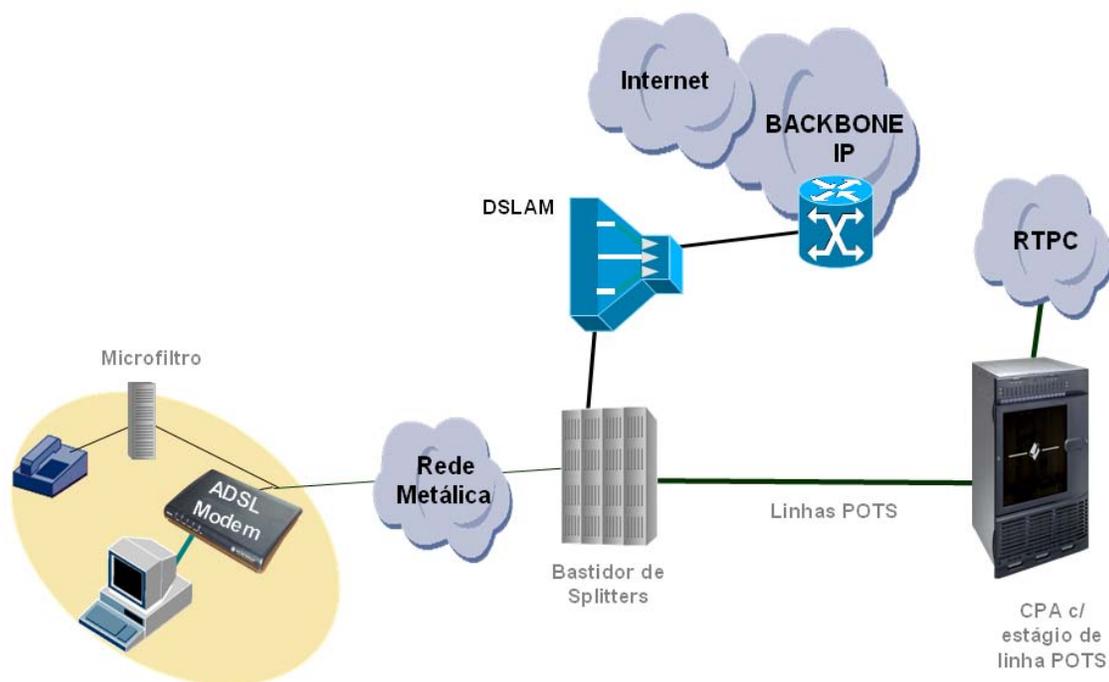


Figura 1.1: Sistema DSL típico, com o modem do usuário à esquerda conectado à rede metálica e o *Central Office* à direita a partir do bastidor de *splitters* conectado à rede metálica.

1.2 Contexto e Motivação

O enlace telefônico, que é a infra-estrutura básica da tecnologia DSL, é um aspecto importante que determina a máxima taxa que pode ser alcançada por um sistema DSL. Na Figura 1.2 está representado genericamente um enlace telefônico com n_s seções, com cada seção constituída de um comprimento l e uma bitola g . Todas as seções são seriais (ou seja, seguem o caminho direto até o usuário), com exceção da terceira seção que é uma derivação à linha principal do assinante chamada também de *bridged-tap*. Como a banda de voz de um sistema telefônico tem uma natureza banda-estreita (dc até 4 kHz), nem todos os enlaces existentes são apropriados para redes DSL, devido às limitações na qualidade do cabo ou fatores limitantes existentes na linha do assinante, como a existência de *bridged-taps* ou o excessivo número de mudanças de bitolas, que poderiam ser previstas através da documentação da rede telefônica. Entretanto a documentação adequada nem sempre é feita durante a expansão da infra-estrutura existente ou em alguns casos foi perdida, impossibilitando aos provedores de serviço prever a viabilidade de um enlace para aplicação em sistemas DSL, assim como a taxa de dados que poderia ser transmitida pela tecnologia nesse enlace. A motivação para um sistema de identificação de topologias de enlaces DSL é justamente a qualificação do enlace (em inglês *Loop Qualification*, ou simplesmente LQ) para determinar as taxas de transmissão que poderiam ser oferecidas ao usuário para um determinado enlace telefônico. Outro aspecto importante seria a possibilidade de expansão das taxas de transmissão de um enlace, por exemplo, através da determinação da posição das *bridged-taps*, que ao serem retiradas proporcionariam um aumento nas taxas de transmissão de dados daquele enlace [3]. Diversos trabalhos de identificação de enlaces, baseados em reflectometria no tempo e na frequência, foram propostos nos últimos anos [9–12].

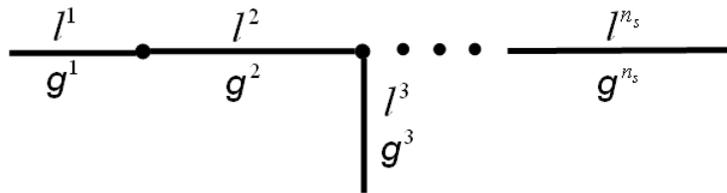


Figura 1.2: Parâmetros do enlace Θ a serem determinados. Nesse exemplo, a terceira seção é uma *bridged-tap* enquanto todas as outras são seriais.

1.2.1 Métodos de Identificação de Linha

Um método de LQ clássico consiste na mineração de bancos de dados existentes com informações das plantas telefônicas para selecionar enlaces apropriados para aplicação DSL [9]. No entanto, muitas vezes esses bancos de dados encontram-se desatualizados ou em alguns casos foram perdidos, tornando esse método pouco confiável ou impraticável. Além disso, métodos de LQ não automáticos apresentam custos elevados e estão mais sujeitos a erros devido à intervenção humana. Outro método alternativo para LQ, que é o foco desta proposta, é a identificação completa da topologia do enlace: o comprimento total do enlace, número de seções, comprimento e o bitola de cada seção do enlace. A identificação de topologia também inclui a identificação do número de *bridged-taps*, assim como suas localizações e seus comprimentos. Um método preciso e confiável de identificação de topologias permite realizar a qualificação de enlace de forma eficaz e também pode fornecer informações para atualizar e corrigir os registros de enlaces existentes que, assim, podem ser acessados para dar suporte às operações de engenharia, instalação e manutenção [11].

Métodos de identificação de topologias podem ser classificados em duas classes: métodos baseados em medidas em uma porta, chamados de métodos de SELT (*Single-Ended Line Testing*) que normalmente são executados a partir da central, e métodos baseados em medidas em duas portas, chamados de DELT (*Dual-Ended Line Testing*), que requerem a comunicação entre o equipamento de medição no CO e o modem na localidade do assinante (referida em inglês como *Customer-Premises*, CP).

Medidas de SELT podem ser baseadas em reflectometria no domínio do tempo (*Time-Domain Reflectometry*, ou simplesmente TDR) ou no domínio da frequência (*Frequency-Domain Reflectometry*, ou FDR) [11][13]. Essencialmente, ambos os métodos enviam um sinal de prova através do enlace e analisam suas reflexões, através do que é comumente chamado reflectograma.

Em [9], os autores apresentam um modelo para analisar os ecos reais e espúrios do reflectograma obtido via TDR, assim como um modo para extrair do reflectograma a parcela do sinal devido ao decaimento lento, decorrente da natureza RLC do enlace, e assim melhorar a detecção de ecos fracos. Mas esse trabalho não apresenta nenhuma metodologia de caracterização dos ecos para identificação das descontinuidades do enlace (mudança de bitolas, *bridged-taps*, entre outras). Em [11], um processo iterativo de compensação (chamado em inglês de *de-embedding*) para o reflectograma TDR é proposto. Nele são identificados os tipos de descontinuidades, usando um modelo de linha e comparações através de MSE (*Mean Squared-Error*), e em seguida as topologias e os comprimentos das seções do enlace, sendo as respostas dessas topologias compensadas do reflectograma atual, resultando em um reflec-

tograma com influência apenas das seções ainda a serem identificadas. Em [14], a técnica de identificação apresentada em [11] é estendida, usando as técnicas de TDR melhoradas descritas em [9] e também um método adicional que usa estatísticas dos enlaces telefônicos para aumentar a precisão do processo de identificação. Essa técnica apresenta bons resultados para enlaces com *bridged-taps*, mas falha na identificação de enlaces com apenas seções seriais, encontrando enlaces com número de seções inferior aos enlaces originais.

Outros métodos de identificação de topologias utilizam medidas do parâmetro de espalhamento $S_{11}(f)$. Em [15], é proposto o uso do parâmetro de espalhamento $S_{11}(f)$ para estimar a função de transferência do enlace quando informação *a priori* da topologia do enlace está disponível. Apesar desse método apresentar bons resultados para enlaces curtos e médios, a consideração sobre o conhecimento parcial ou total de antemão da topologia do enlace limita a aplicabilidade dessa técnica. Além disso, essa técnica pode apresentar resultados fisicamente não realizáveis (valores negativos para os parâmetros do enlace). Outro método de identificação de topologias trabalha a partir da conversão das medidas do parâmetro de espalhamento $S_{11}(f)$ para o domínio do tempo $s_{11}(t)$ e com a identificação das características mais importantes do reflectograma de $s_{11}(t)$, como atraso, amplitude e energia das reflexões. Além disso, é proposto a interpretação automatizada dessas informações por um sistema especialista, baseado em redes *bayesianas*, para estimação da topologia do enlace [12], [16,17]. Essa técnica apresenta a desvantagem de ser complexa em comparação às técnicas descritas anteriormente, com pouca informação sobre sua implementação. Outro aspecto limitante dessa técnica, é a restrição das bitolas que são usadas no processo de estimação, sendo consideradas apenas as bitolas 0,40, 0,50 e 0,63 mm e sempre na ordem crescente. Na prática outros valores de bitola podem ocorrer, como 0,32 e 0,90 mm, e além disso as bitolas podem se repetir ao longo do enlace. Dessa forma, essa técnica apresenta resultados para um número limitado de enlaces, não podendo ser aplicada a outros enlaces que podem ser importantes em termos de ocorrência na planta telefônica.

Outros métodos de identificação de topologias, e, adicionalmente, de detecção de emendas mal conectadas, baseados em FDR, são apresentados em [18,19]. A técnica FDR envolve a transmissão de um sinal de prova faixa larga através do enlace sob teste. A sobreposição dos espectros do sinal transmitido e recebido revela frequências nas quais interferências construtivas e destrutivas ocorrem, gerando ondas estacionárias em frequências mais baixas que as frequências do sinal transmitido. A utilidade do FDR é justamente devido ao fato que a amplitude pico a pico da sobreposição dos espectros é dependente da perda do cabo e da magnitude do descasamento de impedância que ocorre no final do enlace sob teste, enquanto que a frequência de modulação é uma função do comprimento da linha para aquele descasamento de impedância. Sistemas especialistas geralmente são necessários para analisar as formas de

onda de saída de FDR para identificar as possíveis causas do descasamento de impedância, que podem ser: mudanças de bitola, emendas mal conectadas e *bridged-taps* [18–20].

Em geral, os métodos de identificação de topologias previamente descritos necessitam de informações *a priori* sobre o enlace para obter resultados precisos. Essas informações podem ser o comprimento total do enlace, a velocidade de propagação do sinal através do enlace, entre outras; mas em geral, na prática, essas informações *a priori* não estão disponíveis ou não são confiáveis, o que mostra a necessidade de desenvolver métodos que não necessitem dessas informações.

A maioria dos métodos de identificação de topologias existentes na literatura é baseada em medidas de SELT, devido às dificuldades para obter medidas de duas portas baseadas em DELT. Uma única exceção é o método descrito em [21], que detecta e identifica *bridged-taps*, usando medidas de função de transferência. Porém, esse método analisa apenas um tipo de enlace bem específico, com duas seções seriais e uma *bridged-tap*, para estimação do comprimento e da posição dessa *bridged-tap*. As bitolas não são obtidas por esse método. Apesar das limitações explícitas desse método, foram realizados testes para enlaces com mais de uma *bridged-tap*, que mostraram que esse método não é capaz de detectar corretamente a segunda *bridged-tap*, além de alguns casos estimar com grande erro a primeira *bridged-tap*, que pode ser devido a influência dessa segunda *bridged-tap*.

A partir do advento dos padrões G.992.3 e G.992.5 para ADSL 2 e ADSL 2+, as medidas de função de transferência podem ser obtidas diretamente através da nova funcionalidade de diagnóstico de enlaces dos modems ADSL 2 e ADSL 2+ [6, 7]. Esses modems, instalados na casa do assinante, se comunicam com o multiplexador de acesso à linha digital do assinante, o DSLAM, localizado na central e executam medidas de DELT para obter, assim, a função de transferência do enlace. Atualmente é disponibilizado pelos provedores de serviço DSL, um kit de instalação, incluindo um modem ADSL, que ao ser conectado à linha do assinante torna possível executar testes de LQ baseados em DELT sem a necessidade de enviar um técnico à localidade do usuário. Essa nova funcionalidade, baseada em medidas de DELT, impulsiona o desenvolvimento de novas técnicas, mais confiáveis e precisas, para a identificação de topologias e supervisão em sistemas DSL.

1.3 Objetivos do Trabalho

A linha digital de assinante, que conecta o equipamento do usuário ao CO, tem diversos fatores limitantes que não o são para o serviço de telefonia convencional (POST - *plain old telephony service*), apesar de utilizarem o mesmo enlace telefônico. Isso ocorre devido o serviço

DSL explorar uma faixa de frequência mais ampla. Como consequência, os equipamentos de teste de uma rede POTS não são capazes de qualificar a linha de assinante para transmissão DSL. Esses fatores limitantes que causam atenuação excessiva, como *bridged-taps* e o número grande de emendas ao longo do enlace, limitam o alcance do sistema e as taxas que podem ser transmitidas. O desafio é caracterizar completamente um enlace, isto é, identificar sua topologia: comprimento total do enlace, número de seções, comprimento e bitola de cada seção, e número de *bridged-taps* e suas posições e comprimentos. A identificação da topologia do enlace permite às companhias de telefone atualizar e corrigir os registros da sua planta física de cobre, e esses registros podem dessa forma ser acessado para suportar operações de engenharia, provisionamento e manutenção. Desse modo, o desenvolvimento de técnicas de identificação de topologia automática é muito importante para redução do custo durante o estágio de implantação do sistema e mesmo depois, durante testes de monitoramento preventivos contra falhas eminentes de serviço.

O presente trabalho descreve uma metodologia para identificação de topologias que utiliza os novos recursos de medidas baseados nos padrões ITU-T G.992.3 e G.992.5. A metodologia proposta usa medidas de uma e de duas portas, SELT e DELT, respectivamente. Como dito anteriormente, apenas um trabalho é conhecido que trabalha com a função de transferência (medida de DELT), e não para identificação da topologia completa do enlace [21]. Dessa forma, a metodologia proposta é a primeira a utilizar recursos de diagnóstico de enlace para identificação de topologias para sistemas DSL. Para uma visão geral, a Figura 1.3 ilustra o sistema DSL na qual a metodologia proposta pode ser aplicada, utilizando os novos recursos de medição. Basicamente, o modem se comunica com o DSLAM para realizar as medidas de DELT e utiliza o enlace DSL (par trançado) para enviar as medidas que foram realizadas ao DSLAM. Esse por sua vez envia as medidas, que também podem incluir medidas de SELT, para o escritório de supervisão/manutenção através da rede banda larga da empresa (*broadband network*), usando por exemplo, o protocolo SNMP (*Simple Network Management Protocol*). No escritório de supervisão/manutenção, a metodologia proposta é automaticamente executada e é realizada a identificação da topologia do enlace, sendo em seguida aplicados os procedimentos de qualificação ou manutenção ao enlace em análise.

Mais especificamente, a metodologia proposta usa o parâmetro de espalhamento S_{11} e a função de transferência H , no domínio da frequência e no domínio do tempo, para identificar a topologias de enlaces DSL através de um algoritmo genético (AG) multiobjetivo, uma técnica local de otimização e técnicas analíticas. A metodologia proposta foi desenvolvida em parceria com a empresa Ericsson e resultou em um produto, sendo denominada de LQ-SUITE (Loop Qualification SUITE). O LQ-SUITE é constituído de três módulos: o TIMEC (*Topology Identification via Model-based Evolutionary Computation*) que é um método que utiliza

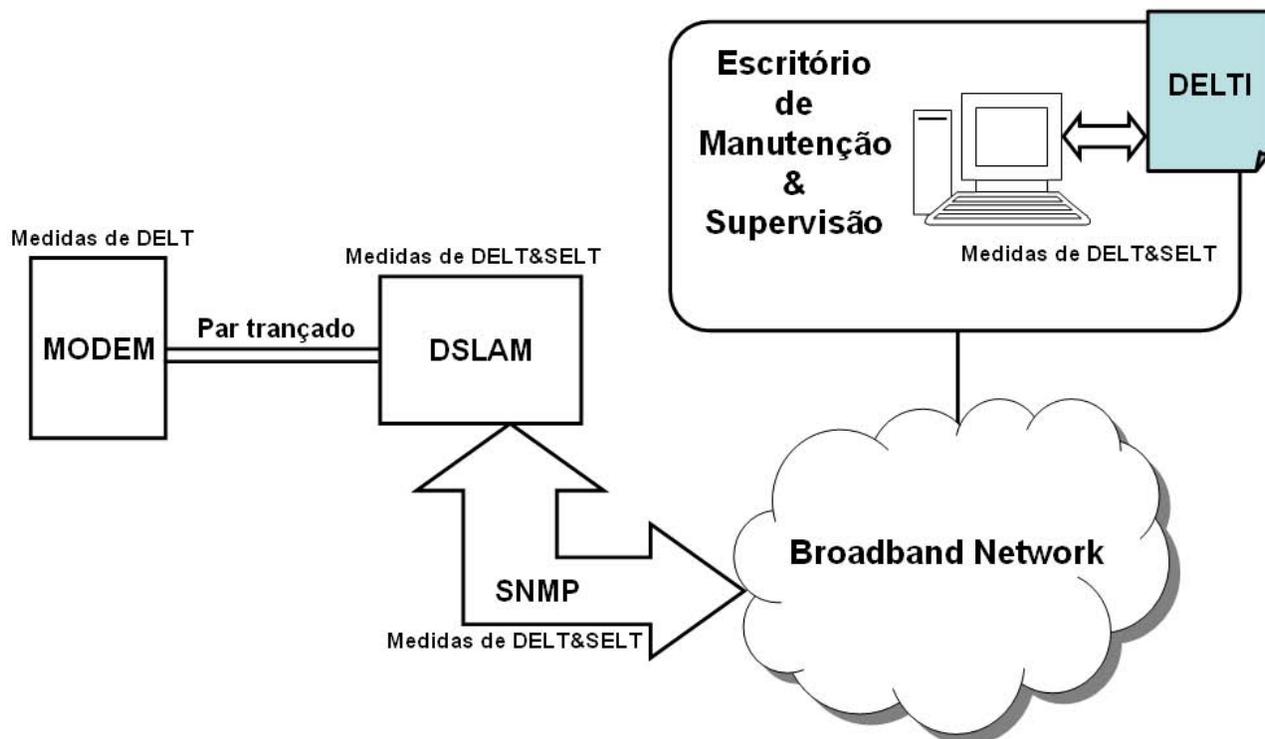


Figura 1.3: A metodologia DELTI aplicada em um sistema DSL.

modelos de linha e duas funções custos, uma relacionada a medidas de uma porta (parâmetro de espalhamento S_{11} ou impedância de entrada Z_{in}) e outra relacionada a medidas de duas portas (função de transferência H), para avaliar as soluções geradas pelo AG, e assim aplicar o procedimento de seleção por ordenamento em frentes de Pareto, baseado na técnica NSGA-II (*Non-Dominated Sorting Genetic Algorithm- II*). O produto consiste de mais dois módulos, que são utilizados para obter informação *a priori* sobre o enlace. O primeiro deles baseado na análise da resposta impulsiva do enlace chamado de IRA (*Impulse Response Analysis*) e o segundo baseado na análise da resposta no tempo de medidas de uma porta chamado de FSF (*First Section Finder*).

1.4 Contribuições da Tese

A principal contribuição da tese é estimação precisa da topologia do enlace que pode ser usada para atualizar o banco de dado de enlaces da empresa provedora do serviço DSL, que poderá gerenciá-los mais eficazmente e de forma otimizada. Dessa forma, a empresa poderá dizer rapidamente se um serviço DSL pode ser aplicado a um dado enlace. Além disso, com esses dados também é possível estimar precisamente a capacidade do enlace, evitando

desperdício de recursos, com a subestimação da taxa do enlace, e também a insatisfação do usuário, com uma estimativa da taxa acima do que o enlace realmente pode oferecer. É possível também estender a capacidade do enlace, com a identificação de *bridged-taps* e a sua retirada, com um custo mais baixo, pois a posição das *bridged-taps* pode ser encontrada mais facilmente, já que foi previamente identificada pelo LQ-SUITE. Essa metodologia foi comparada com uma técnica recente de identificação de enlaces baseada em TDR e apresentou desempenho melhor que essa técnica. Além disso, é uma técnica mais geral que outras técnicas descritas na literatura, como a técnica dada em [17], que utiliza um conjunto mais restrito de bitolas e sempre considerando que as bitolas ocorrem em ordem crescente. Como na prática as bitolas podem se repetir e ocorrer na ordem inversa (com as bitolas maiores antes das menores) ao longo do enlace, essa técnica apresenta limitação em sua aplicabilidade.

Além dessas contribuições mais importantes, serão descritas outras contribuições, que são aspectos chave relacionados à precisão, confiabilidade e eficiência computacional da metodologia. Como a técnica de controle adaptativo dos parâmetros do AG que modifica os valores das probabilidades de cruzamento e mutação, além do número de gerações (iterações) do AG, conforme a velocidade de variação da aptidão das soluções do AG. Essa técnica possibilita, baseada na eficiência do AG em encontrar melhores soluções a cada geração, uma pesquisa mais ampla do AG no início do processo de otimização, evitando sua convergência prematura, mas tende a realizar uma busca mais restrita ao final do processo de otimização para um refinamento maior da solução encontrada. Outra contribuição são as técnicas de decodificação, que realizam a **convensão** dos parâmetros do AG para o domínio do problema em otimização. Essas técnicas restringem o espaço de busca, baseado no conhecimento da planta telefônica, evitando que o método de otimização pesquise soluções que não são factíveis ou inviáveis do ponto de vista prático. Isso aumenta a eficiência da metodologia e diminui o seu tempo de processamento. Essas técnicas de decodificação são aplicadas ao método híbrido de otimização (ou seja, ao AG e a técnica local de otimização) que é outra contribuição da tese. Baseado nas características do processo de identificação de enlaces DSL, é utilizada uma técnica de otimização local que interage com o AG ao longo do processo de otimização, aumentando sua precisão e reduzindo o seu tempo de otimização. Além disso, com auxílio desse método híbrido de otimização, o enlace é iterativamente reconstituído através de um processo denominado de realimentação de seções, que se utiliza da forte influência da medida de uma porta na estimativa das primeiras seções do enlace. Esse processo aumenta a precisão da identificação do enlace e reduz as chances de convergência prematura do AG. Por fim, considerando a importância da utilização de medidas confiáveis, foi desenvolvida em co-autoria uma ferramenta automática de medição, que realiza os procedimentos de controle dos equipamentos de medição e coleta dos dados quase sem interferência do usuário, diminuindo assim

as chances de ocorrerem erros no processo de medição; além disso, realiza a análise estatística, a detecção e eliminação de discrepâncias (chamados comumente de *outliers*) das medições obtidas, obtendo um conjunto de medições mais preciso a ser utilizado pela metodologia de identificação de topologias.

Todos esses itens podem ser resumidos no diagrama de blocos mostrado na Figura 1.4. Seguindo o fluxo de dados: primeiramente, após as medições, é realizado o tratamento estatístico dos dados medidos com a detecção e eliminação de discrepâncias, em seguida a extração de informações *a priori* sobre o enlace pelas técnicas analíticas de pré-processamento IRA e FSF e por fim a aplicação de TIMEC, que é baseado em modelos de linha, otimização evolucionária multiobjetivo e otimização local. Nesse módulo são usadas as técnicas de codificação\decodificação e o controle auto-adaptativo dos parâmetros do AG. Assim, o LQ-SUITE, constituído por todos esses módulos, estima a topologia completa do enlace, que é o objetivo final da tese.

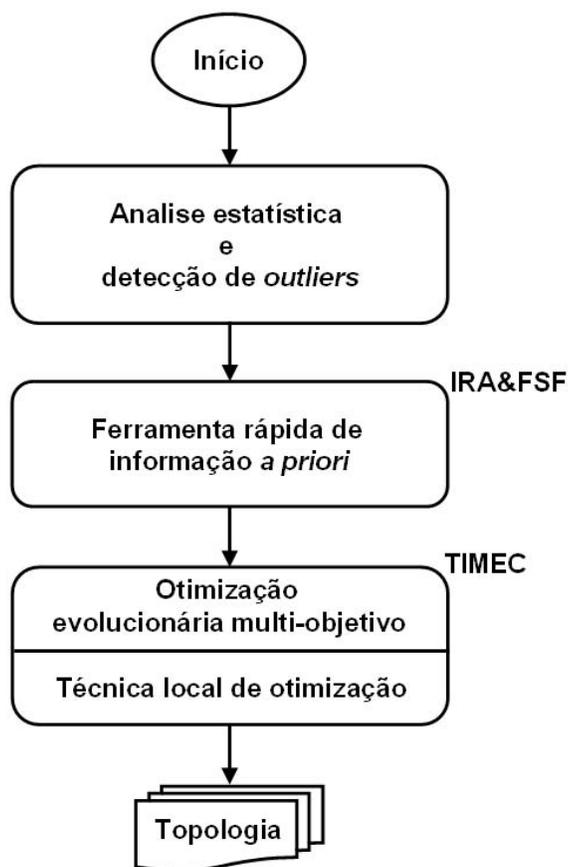


Figura 1.4: Visão resumida das várias áreas de ação e possíveis contribuições associadas a esta proposta de tese.

1.5 Estrutura do Trabalho

No capítulo 1 foram discutidos o estado da arte e as motivações para o desenvolvimento desse trabalho. O restante do trabalho está organizado da seguinte maneira:

- Capítulo 2 - são discutidos aspectos relacionados a análise e modelagem do par trançado, que é meio de comunicação de sistemas DSL;
- Capítulo 3 - trata dos aspectos relacionados ao algoritmo genético, especialmente o algoritmo genético multiobjetivo, sendo apresentada antes a teoria básica sobre otimização multiobjetivo. Por fim, é descrito o algoritmo genético multiobjetivo NSGA-II em que este trabalho se baseia;
- Capítulo 4 - nesse capítulo é discutido o método TIMEC que usa algoritmos genéticos multiobjetivos e técnicas locais de otimização. O TIMEC é a principal parte da metodologia de identificação de enlaces LQ-SUITE. Além disso, são mostrados as especializações do processo de otimização, como a estratégia de controle adaptativo dos parâmetros do AG e o processo de decodificação. É também descrita a ferramenta de medição automática, que realiza a coleta e tratamento dos dados de medição;
- Capítulo 5 - nesse capítulo são apresentadas as ferramentas de pré-processamento. Primeiramente, é descrita a técnica que realiza a análise da resposta impulsiva do enlace denominada de IRA. A seguir o FSF, que usa medidas de TDR para obter informações sobre o enlace;
- Capítulo 6 - nesse capítulo são apresentados os resultados obtidos, primeiro de cada módulo trabalhando separadamente e a seguir os resultados de todos os módulos atuando em conjunto. Para avaliação da metodologia foram considerando aspectos relacionados ao custo computacional e a precisão das respostas obtidas;
- Capítulo 7 - as conclusões e propostas para prosseguimento do trabalho são apresentadas.

Capítulo 2

Análise e Modelagem do Par Trançado

Este capítulo trata dos principais tópicos relacionados à análise e modelagem do par trançado, meio de comunicação de sistemas DSL. A primeira seção trata da análise do par trançado através da teoria padrão ABCD. A segunda seção trata de aspectos relacionados aos modelos de linha utilizados para obter os parâmetros de interesse de enlaces DSL. Na última seção são apresentados os comentários finais.

2.1 Análise do par trançado

O cabo de par trançado é simplesmente composto de dois fios trançados entre si, alimentados eletricamente usando o modo diferencial (no qual as correntes fluem em um sentido em um fio e retornam em sentido contrário pelo segundo fio criando uma diferença de potencial entre eles) para reduzir os efeitos de interferências externas, como por exemplo o *crosstalk* que ocorre devido a influência de outros cabos de par trançado [22]. Como os fios do par trançado são fisicamente próximos e são simétricos entre si, qualquer influência externa em um dos fios afetará o outro de forma semelhante, e assim, a interferência será extraída quando for usado o modo diferencial para obter o sinal de informação.

Nesta seção serão discutidos aspectos relacionados às características do par trançado e à sua análise.

2.1.1 Características do par trançado

Tipos diferentes de cabos de par trançado, que também apresentam (mesmo para tipos iguais) níveis de qualidade diferentes, constituem a infra-estrutura de *cabeamento* existente.

Cabos de par trançado normalmente são constituídos por 25 a 100 pares trançados, que são subdivididos dentro do cabo em feixes de 3 a 4 pares. Existem dois tipos de cabos de par trançado:

- Unshielded Twisted Pair - UTP (par trançado sem blindagem): São os cabos de rede de par trançado sem blindagem. Esse tipo de cabo é muito popular, justamente por ser bastante fácil de se trabalhar e muito barato. Existem várias categorias de cabos UTP, que indicam a qualidade do cabo e as aplicações em que podem ser usados. Os mais primitivos são os cabos categoria 1, que eram usados em instalações telefônicas antigas. Atualmente, os mais usados são os cabos categoria 5, que podem ser usados em redes Ethernet de 10 ou 100 megabits.
- Shield Twisted Pair - STP (cabo com blindagem): Cabos de rede blindados, usados em ambientes com fortes interferências eletromagnéticas. Eles contam com uma proteção extra se comparado aos cabos UTP. Além do entrelaçamento dos fios, eles possuem uma blindagem externa. Podem ser encontrados com blindagem simples ou com blindagem par a par. São mais caros, necessitam de aterramento e tem menor utilização.

A taxa de torção entre os pares trançados de um feixe varia em relação aos outros feixes em um mesmo cabo para uma maior redução do *crosstalk* dentro do cabo. A taxa de torção normalmente varia de 12 a 40 voltas por metro [3]. Cada fio também é envolvido com algum tipo de material isolante feito de papel ou de material plástico. Além do material de isolamento, os cabos de par trançado são caracterizados também pelo diâmetro do fio de cobre, ou bitola, do cabo. Para o padrão ETSI (*European Telecommunications Standards Institute*) as bitolas podem variar de 0,32 mm a 0,90 mm. Para o padrão americano, é usada a designação AWG (*American Wire Gauge*) com valores de bitolas de 19 AWG (0,912 mm) a 26 AWG (0,404 mm).

As características elétricas dos cabos de par trançado são definidas usando o modelo clássico de linha de transmissão. Esse modelo incorpora um conjunto de quatro parâmetros distribuídos, que incluem a indutância e resistência série e a capacitância e condutância paralela, também conhecidos como parâmetros RLCG do cabo [23]. A indutância série representa a auto-indutância total dos dois condutores e a capacitância em paralelo ocorre devido à proximidade dos dois condutores. Os parâmetros RLCG são também conhecidos como parâmetros primários da linha de transmissão e são funções da frequência. A Figura 2.1 ilustra uma representação gráfica de um segmento de par trançado para um comprimento de linha Δz .

A teoria clássica de linhas de transmissão permite descrever o comportamento de tensão e corrente, para dependência temporal harmônica, a partir do conjunto de equações:

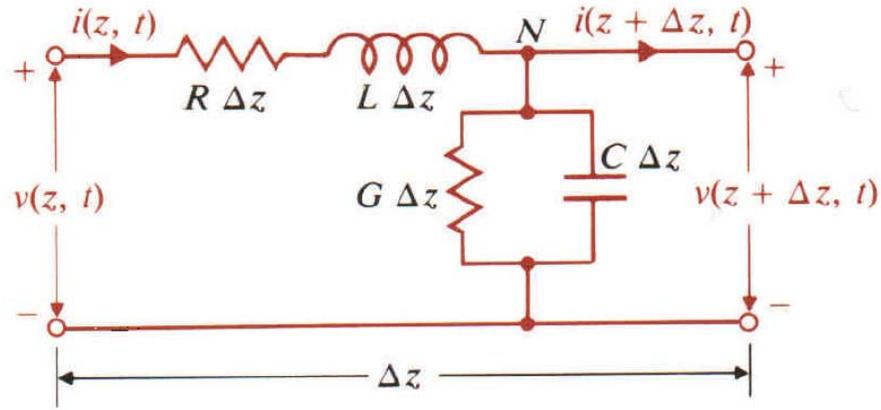


Figura 2.1: Modelo de linha de transmissão por unidade de comprimento, mostrando a indutância L , a resistência R , a capacitância C e a condutância G .

$$\frac{d^2 V_s}{dz^2} - \gamma^2 V_s = 0 \quad (2.1)$$

e

$$\frac{d^2 I_s}{dz^2} - \gamma^2 I_s = 0 \quad (2.2)$$

onde

$$\gamma(f) = \alpha(f) + j\beta(f) = \sqrt{[R(f) + j\omega L(f)][G(f) + j\omega C(f)]} \quad (2.3)$$

sendo γ a constante de propagação (dado em 1/m), e as suas partes reais e imaginárias são a constante de atenuação, $\alpha(f)$ (Np/m), e a constante de fase, $\beta(f)$ (rad/m), respectivamente.

Conforme a definição de impedância característica Z_0 , que é a razão entre a onda de tensão e a onda de corrente, que se propagam no sentido positivo, em qualquer ponto da linha, a impedância característica Z_0 é descrita por:

$$Z_0(f) = \sqrt{\frac{R(f) + j\omega L(f)}{G(f) + j\omega C(f)}} \quad (2.4)$$

A constante de propagação γ e a impedância característica Z_0 , também chamados de parâmetros secundários da linha, são propriedades importantes da linha porque ambas dependem dos parâmetros de linha R, L, C e G e da frequência de operação. O recíproco de Z_0 é a admitância característica Y_0 , isto é, $Y_0 = 1/Z_0$. Os parâmetros secundários da linha são usados para modelar os cabos de par trançado quando são processadas a síntese e análise dos enlaces dos assinantes.

2.1.2 Teoria de Matriz ABCD

Itens como o comprimento l , a bitola g e o tipo de cada seção, assim como o número de seções do enlace definem a topologia do enlace de um assinante DSL, como mostrado na Figura 2.2. O cabo de par trançado pode ser modelado como uma rede de duas portas (quadripolo) mostrada na Figura 2.3, e assim pode ser representada pela matriz na Equação 2.5, chamada de matriz ABCD [24,25], permitindo que importantes características do enlace do assinante possam ser obtidas, tais como impedância de entrada Z_{IN} , parâmetro de espalhamento S_{11} , função de transferência H e perda por inserção I_L a partir de sua topologia.

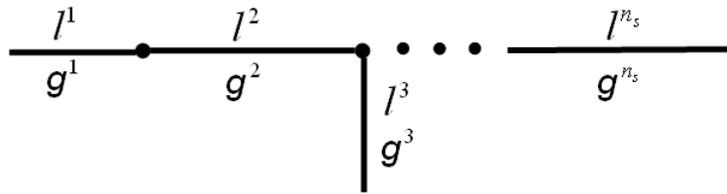


Figura 2.2: Representação de um enlace DSL. Nesse enlace, a terceira seção é uma *bridged-tap* enquanto as outras seções são seriadas. O número total de seções é n_s .

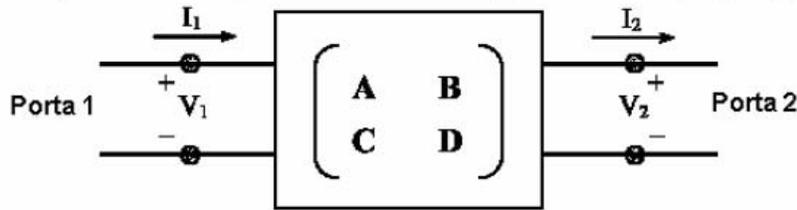


Figura 2.3: Quadripolo ABCD.

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_2 \\ I_2 \end{bmatrix} \quad (2.5)$$

A Equação (2.6) descreve a matriz ABCD para o segmento serial de um enlace, que relaciona o comportamento da impedância característica da linha de transmissão e da constante de propagação ao longo do comprimento. A derivação (*bridged-tap*), presente na linha, é definida pela matriz ABCD, como na Equação (2.7).

$$\mathbf{T}_s = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \cosh(\gamma l) & Z_0 \sinh(\gamma l) \\ \frac{1}{Z_0} \sinh(\gamma l) & \cosh(\gamma l) \end{bmatrix}, \quad (2.6)$$

$$\mathbf{T}_{bt} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1}{Z_0} \tanh(\gamma l) & 1 \end{bmatrix}, \quad (2.7)$$

onde l é o comprimento da seção correspondente. Note que γ e Z_0 são valores complexos dependentes da frequência, mas por simplicidade esta dependência é omitida aqui. Os valores nominais para γ e Z_0 são obtidos através de modelos de linha, considerando-se a bitola de cada seção de linha e o tipo de material do qual o cabo é constituído [26].

Para um enlace composto por seções de bitolas e/ou comprimentos diferentes (cascata de várias seções), pode-se multiplicar as matrizes ABCD de cada seção do enlace para obter a matriz ABCD global. Assim, se cada seção de linha s é representada por sua matriz de transmissão \mathbf{T}^s , a matriz de transmissão direta global \mathbf{T}_f , que representa o enlace do assinante inteiro, é determinada por [2]:

$$\mathbf{T}_f = \mathbf{T}^1 \times \mathbf{T}^2 \dots \times \mathbf{T}^{m_s}. \quad (2.8)$$

Para análise de um enlace, são necessárias informações adicionais além da matriz ABCD dos segmentos de par trançado. Essa informação inclui a impedância da fonte Z_S , que é normalmente a impedância do transmissor de dados usado para enviar as informações, ou a impedância do equipamento de medida, se um processo de análise do enlace de assinante está em andamento; e também, a impedância de terminação ou de carga, Z_L , do cabo de distribuição, que será assumida em circuito aberto, para medidas de uma porta (SELT), e casada para medidas de duas portas (DELT).

A impedância de entrada do enlace de assinante é definida como a relação da tensão e da corrente na porta de entrada do enlace. Em termos de parâmetros ABCD, a impedância de entrada é dada por [2]:

$$Z_{IN} = \frac{V_1}{I_1} = \frac{AZ_L + B}{CZ_L + D}. \quad (2.9)$$

O parâmetro de espalhamento S_{11} é definido como a razão entre a onda de saída (refletida) e a onda de entrada na mesma porta (Apêndice B), em termos de parâmetros ABCD, é dado por [15]:

$$S_{11} = \frac{AZ_L + B - CZ_S Z_L - DZ_S}{AZ_L + B + CZ_S Z_L + DZ_S}. \quad (2.10)$$

A função de transferência do enlace de assinante é definida como a relação entre a tensão através da porta de saída e a tensão através da porta de entrada. Em termos de parâmetros ABCD, a função de transferência é dada por [2]:

$$H = \frac{Z_L}{AZ_L + B + CZ_S Z_L + DZ_S}. \quad (2.11)$$

A perda de inserção do enlace é definida como a relação da potência entregue à carga através do enlace e da potência entregue à carga com o enlace não estando presente. É basicamente a perda que é inserida entre o CO e o modem do assinante devido ao enlace e é dado em decibéis como [2]:

$$I_L = 10 \log \left(\left| \frac{P_{nolloop}}{P_{loop}} \right| \right) = 20 \log \left(\left| \frac{AZ_L + B + CZ_S Z_L + DZ_S}{Z_L + Z_S} \right| \right). \quad (2.12)$$

2.2 Modelos de linha

Os modelos devem ser capazes de prever o quanto eficientemente uma fonte excita a linha, como os sinais se propagam ao longo da linha e como o sinal chega ao usuário. A criação de um modelo realístico que capture com precisão a condição real de uma linha é extremamente útil para o refinamento de projetos de sistemas ADSL e VDSL. Vários modelos para canal de comunicação foram propostos; cada um assume certo número de simplificações e restrições que podem reduzir a precisão do modelo quando aplicado às linhas telefônicas reais [27]. Os sistemas DSLs são relativamente simples para projeto de sistemas que trabalham em simulação e em testes laboratoriais de poucas especificidades, mas muito difíceis para operar em capacidades máximas de uso quando sujeitos ao ambiente hostil da rede real [28]. Esses modelos basicamente fornecem os valores nominais para os parâmetros primários e/ou secundários.

Basicamente, dois aspectos devem ser levados em consideração ao escolher um modelo de linha:

- O modelo deve estar em acordo com as relações de causalidade de Hilbert;
- A quantidade de parâmetros que constituem o modelo.

Quando a parte real e a parte imaginária da impedância série e admitância paralela estiverem de acordo com a relação de Hilbert, o modelo de linha é causal, isto é, suas repostas chegam depois da excitação. Portanto, a resposta impulsiva é unilateral: $h(t) = 0$ para $t < 0$, e

a função de transferência do sistema $H(\omega)$ é uma função analítica no plano complexo. A parte real e imaginária da função analítica $H(j\omega)$ forma um par da transformada de Hilbert [29]:

$$A(\omega) = \frac{-1}{\pi} P \int_{-\infty}^{+\infty} \frac{B(\lambda)}{\lambda - \omega} d\lambda \quad (2.13)$$

$$B(\omega) = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{A(\lambda)}{\lambda - \omega} d\lambda \quad (2.14)$$

onde P é o valor principal de Cauchy. $A(\omega)$ e $B(\omega)$ são unicamente relacionados através das relações (2.13) e (2.14). Alguns modelos de linha conhecidos e que estão nos padrões de sistemas DSL estavam focados apenas no domínio da frequência e, por isso, apresentam comportamento não realísticos no domínio do tempo, como a informação viajar mais rápido que a luz e a não causalidade [30].

O número de parâmetros que constituem o modelo é um importante aspecto relacionado a habilidade do modelo de se ajustar aos dados medidos. Geralmente são utilizadas técnicas de otimização para o ajustar a resposta do modelo às medidas [30], e assim um número grande de parâmetros aumenta a flexibilidade e facilita esse processo de ajuste. Por outro lado, quanto maior o número de parâmetros maior o custo computacional e mais demorado pode ser o processo de ajuste. Outro aspecto importante é a maior dificuldade de encontrar estimativas iniciais dos parâmetros do modelo a serem fornecidas às técnicas de otimização. Por exemplo, os modelos da British Telecom, BT0 e BT1, tem 10 e 13 parâmetros, respectivamente. Esses modelos demandam maior tempo de processamento para realizar o ajuste aos dados medidos. Por outro lado, apresentam uma maior grau de liberdade que por exemplo o modelo VUB que possui apenas 5 parâmetros. No entanto, ambos os modelos da British Telecom são não-causais enquanto o modelo VUB é causal. Muitos modelos de linha no padrão de sistemas DSL não são adequados para identificação da rede com topologia muito complexa devido ao seu excesso de parâmetros [26].

Na literatura são descritos pelo menos dois grupos de modelos de linha [30]. Primeiro, os modelos de linha analíticos, ou físicos, que são baseados na geometria física e propriedades dos materiais do cabo. Facilitam a investigação do comportamento elétrico quando um parâmetro de projeto, por exemplo, o diâmetro do fio, é modificado, embora as respostas medidas não correspondam exatamente aos modelos, devido às variações mecânicas e elétricas da linha, tais como: torção dos fios, múltiplos isoladores, fios vizinhos, entre outras. Um exemplo bem conhecido de modelo físico é o modelo de Kelvin para a impedância série $Z_s(\omega)$ de uma linha de 2 fios com condutores cilíndricos levando em consideração o efeito pelicular. A partir desse modelo podem ser obtidas fórmulas simples para as características elétricas da linha na frequência, mas que não modelam corretamente sua resposta impulsiva [30]. O

segundo grupo de modelos de linha, são os modelos semi-empíricos, ou heurísticos, baseados no comportamento físico e em equações heurísticas obtidas através de medições do cabo. Esses modelos são descritos nos padrões [26][31,32]. Os modelos semi-empíricos podem exibir algumas inconsistências como as descritas no parágrafo anterior.

A seguir são mostrados alguns exemplos de modelos de linha analíticos e heurísticos. A maioria desses modelos tenta obter a impedância série $Z_s(\omega)$ e a admitância paralela $Y_p(\omega)$, que podem ser representadas conforme a Figura 2.4.

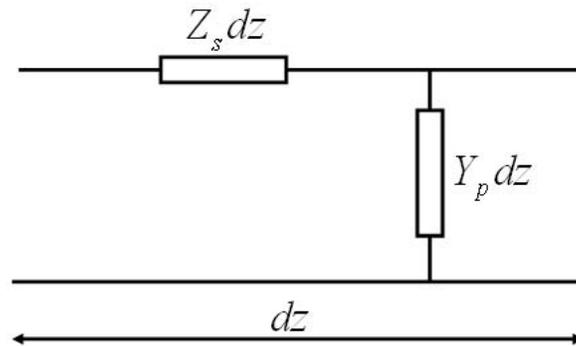


Figura 2.4: A seção de linha infinitesimal.

2.2.1 Modelo RLGC analítico

Neste modelo apenas a resistência R é função da frequência, e os demais são constantes [24]:

- Resistência: $R = R_0 f^{\frac{1}{2}}$
- Indutância: $L = L_0$
- Capacitância: $C = C_0$
- Condutância: $G = 0$

Os valores das constantes são dados na Tabela 2.1 para as bitolas 24 e 26 AWG:

2.2.2 Modelo RLCG pela contribuição ANSI T1E1.4

Um dos modelos mais populares é o modelo empírico British Telecom, ou BT0, focado no modelagem de $\{Z_s, Y_p\}$, sendo usado e descrito em todos os padrões de DSL. Esse modelo é

Tabela 2.1: Parâmetros para o modelo analítico do par trançado

Parâmetros	24 AWG	26 AWG
R_0 (Ω/kft)	0,15	0,195
L_0 (mH/kft)	0,188	0,205
C_0 (nF/kft)	15,7	15,7

capaz de prever as características do cabo de DC a dezenas de MHz. A impedância série Z_s e a admitância paralela Y_p dadas pelo modelo BT0 são descritas nas Equações (2.15) e (2.16):

$$Z_s(f) = \sqrt[4]{R_{oc}^4 + a_c f^2} + j\omega \cdot \left(\frac{L_0 + L_\infty (f/f_m)^{Nb}}{1 + (f/f_m)^{Nb}} \right) \quad (2.15)$$

$$Y_p = (g_0 f^{Nge}) + j\omega \cdot C_\infty \quad (2.16)$$

onde R_{oc} é a resistência DC, a_c é a constante que controla o fator \sqrt{f} , L_0 é a indutância para baixa frequência, L_∞ é a indutância para alta frequência, f_m e Nb são parâmetros que controlam a transição entre os dois parâmetros de indutância, g_0 e Nge permitem adequar a medida de condutância, e C_∞ é o valor de capacitância constante, que assume a permissividade relativa do dielétrico, que envolve o condutor, constante sobre a faixa de frequência de interesse. Nesse modelo, a impedância série é não causal. Isto ocorre devido à resistividade (parte real de Z_s) e a reatância (parte imaginária de Z_s) terem sido modeladas separadamente sem levar em consideração as condições de causalidade de Hilbert. O comportamento não causal desse modelo foi demonstrado claramente em [33]. A Tabela 2.2 ilustra os parâmetros para o par trançado de 24 e 26 AWG.

2.2.3 Modelo RLCG pela contribuição ETSI

O modelo de referência de cabo para simulação em acesso de rede metálica é tratado pelo documento ETSI TM6, sendo aplicado para modelagem na faixa de frequência de sistemas VDSL. Esse modelo, denominado BT1, está descrito a seguir [34]:

$$Z_0 = \left(\frac{1}{\sqrt[4]{R_{oc}^4 + a_c f^2}} + \frac{1}{\sqrt[4]{R_{os}^4 + a_s f^2}} \right)^{-1} + j\omega \cdot \left(\frac{L_0 + L_\infty (f/f_m)^{Nb}}{1 + (f/f_m)^{Nb}} \right) \quad (2.17)$$

$$\gamma = g_0 f^{Nge} + j\omega \left(C_\infty + \frac{C_0}{f^{Nge}} \right) \quad (2.18)$$

Tabela 2.2: Parâmetros para o modelo analítico do par trançado

Parâmetros	24 AWG	26 AWG
R_{oc} (Ω/kft)	174,55888	286,17578
a_c (mH/kft)	0,053073481	0,14769620
L_0 (nF/kft)	$6,1729593 \times 10^{-6}$	$675,36888 \times 10^{-6}$
L_∞ (Ω/kft)	$478,97099 \times 10^{-6}$	$488,95186 \times 10^{-6}$
f_m (Ω/kft)	553720,63	806338,63
Nb (Ω/kft)	1,1529766	0,92930728
g_0 (Ω/kft)	$0,23487476 \times 10^{-12}$	$4,3 \times 10^{-8}$
Nge (Ω/kft)	1,38	0,70
C_∞ (Ω/kft)	50×10^{-9}	49×10^{-9}

onde R_{os} é a segunda componente da resistência DC, a_s é a constante que controla o fator \sqrt{f} e C_0 é capacitância em baixa frequência. O modelo BT1 é também não causal pelo mesmos motivos apresentados na subseção anterior.

2.2.4 Modelo RLCG pela contribuição MAR

O modelo empírico Marconi, ou MAR, ao contrário do modelo BT0 e BT1, respeita a condição de Hilbert entre as componentes real e imaginária de Z_s e Y_p , possibilitando que esse modelo tenha uma resposta impulsiva real. Esse modelo é descrito em [26]. O modelo MAR para Z_s e Y_p é descrito pelas Equações (2.19) e (2.20). Esse modelo é citado na caracterização de cabos de cobre de 50 pares em [35].

$$Z_s = j\omega \cdot L_\infty + R_0 \left(0,25 + 0,75 \sqrt{1 + ar(r+b)/(r+c)} \right) \quad (2.19)$$

$$Y_p = \omega C_f (j + \tan \vartheta) = j\omega \cdot C_{1MHz} \times (jf_{MHz})^{-2\vartheta/\pi} \quad (2.20)$$

onde

$$r = \frac{\mu_0 j f}{0,75^2 R_0} = \frac{j f}{447,6 R_0} \quad (2.21)$$

e

$$C_f = \frac{C_{1MHz}}{f_{MHz}^{\frac{2\vartheta}{\pi}}} \quad (2.22)$$

onde $\mu_0 = 4\pi 10^{-4} H/km$, R_0 é a resistência em corrente contínua por km, L_∞ é a indutância em alta frequência por km, ∂ é o ângulo da capacidade da perda paralela (constante), a é o fator de proximidade, b e c são coeficientes da forma do efeito pelicular (*skin*), e C_{1MHz} é a capacitância por km a 1MHz.

O modelo MAR simplificado (MAR2), usa valores fixos: $b = 2$ e $c = 2,765$ [34].

2.2.5 Modelo VUB

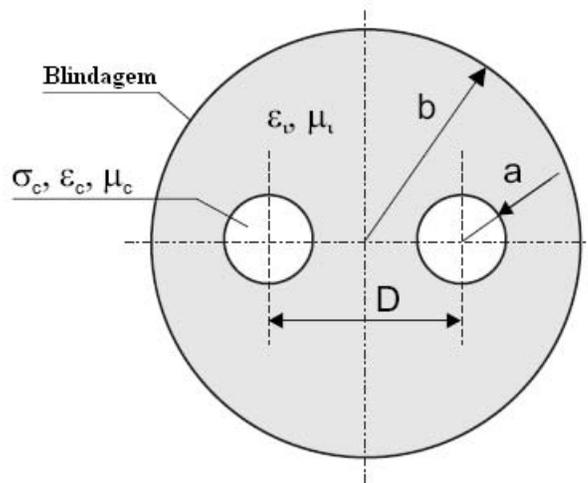


Figura 2.5: Seção transversal de uma linha de 2 fios paralelos.

O modelo VUB é um modelo analítico que descreve a impedância série Z_s e admitância paralela Y_p para uma linha de 2 condutores cilíndricos paralelos de condutividade σ_c , permissividade ε_c e permeabilidade μ_c , conforme mostrado na Figura 2.5. Os condutores são separados por um material isolante de permissividade ε_l e permeabilidade μ_l . Se a linha não possuir blindagem, o valor de b mostrado na Figura 2.5 é igual a infinito ($b = \infty$). A impedância série Z_s , calculada a partir desse modelo, para uma linha de comprimento l com blindagem perfeitamente condutora, é dada por [29, 30][36]:

$$Z_s(\omega)l = a_4 s + a_1 \sqrt{-s} \frac{J_0(a_3 \sqrt{-s})}{J_1(a_3 \sqrt{-s})} + a_1 a_2 a_3 s \Psi \quad (2.23)$$

$$\Psi = \frac{S_1(u)}{a_2 (1 + S_2(u)) + \frac{J_0(a_3 \sqrt{-s})}{J_2(a_3 \sqrt{-s})}} \quad (2.24)$$

$$S_1(u) = 1 - \frac{4u}{1 - u^2} \quad (2.25)$$

$$S_2(u) = \frac{8u(1+u^2)}{(1-u)^2(1+u)^2} \quad (2.26)$$

com

$$a_1 = R_0 \cdot a\sqrt{\mu\sigma} \cdot l \quad (2.27)$$

$$a_2 = \left(\frac{a}{D}\right)^2 \quad (2.28)$$

$$a_3 = a\sqrt{\mu\sigma} \quad (2.29)$$

$$a_4 = L \cdot l \quad (2.30)$$

$$u = \left(\frac{D}{2b}\right)^2 \quad (2.31)$$

onde $s = j\omega$, R_0 é a resistência DC de ambos lados da linha, a é o raio dos condutores, D é a distância entre os centros dos condutores e $J_i(x)$ é a função de Bessel complexa de primeiro tipo de ordem i . Na ausência de blindagem, $b = \infty$ e assim $u = 0$, $S_1 = 1$ e $S_2 = 0$. A indutância externa L , devido aos campos entre os condutores, é dada por [30]:

$$L = \frac{\mu}{\pi} \left[\ln\left(\frac{D}{a}\right) - \ln\left(\frac{1+u}{1-u}\right) \right] \quad (2.32)$$

Se o efeito de proximidade for considerado nulo ($\Psi = 0$), o modelo VUB torna-se o modelo de Kelvin. Para dielétricos de pequenas perdas, o modelo para admitância paralela por unidade de comprimento pode ser dado por [30]: $Y_p = G + j\omega C$ com as constante C e G representando a capacitância e a condutância por unidade de comprimento, respectivamente.

O modelo VUB é causal e descreve com boa precisão diferentes tipos de cabos, conforme descrito em [30].

2.3 Conclusões do Capítulo

Nesse capítulo foram apresentados os conceitos fundamentais sobre a análise e modelagem de cabos de par trançado, mostrando-se como são obtidas as curvas de referências para o processo de otimização, tais como função de transferência e parâmetro de espalhamento. Dos modelos de linha apresentados, deve-se destacar o modelo VUB por apresentar um conjunto reduzido de parâmetros, precisão aceitável e ser causal. Esse será o modelo adotado para implementação da metodologia de identificação de topologias a ser descrita mais a frente. Utilizando o modelo de linha, as quantidades de interesse, como função de transferência (H) e parâmetro de espalhamento (S_{11}), são calculadas para cada indivíduo da população do algoritmo genético e são comparadas as curvas medidas correspondentes para se obter o grau de qualidade do indivíduo, que vai determinar se esse indivíduo será selecionado para gerações futuras do AG.

No próximo capítulo será descrito o processo de otimização multiobjetivo baseado em AG, que conjuntamente com as técnicas de análise do par trançado aqui descrita compõem a metodologia de identificação de enlaces DSL.

Capítulo 3

Algoritmo Genético Multiobjetivo

Nesse capítulo serão apresentados conceitos básicos sobre algoritmos genéticos, especialmente algoritmos genéticos multiobjetivos. A primeira seção descreve os conceitos básicos sobre algoritmos genéticos e seus operadores. A segunda seção descreve os conceitos necessários para entender os aspectos relacionados à otimização multiobjetivo. Na terceira seção são descritos os algoritmos evolucionários multiobjetivos, especialmente o algoritmo genético multiobjetivo NSGA-II, em que esse trabalho se baseia. Finalmente, na quarta seção, são resumidas as principais conclusões deste capítulo.

3.1 Algoritmos Genéticos

3.1.1 Conceito Básico

Os algoritmos genéticos (AGs) são algoritmos de busca e otimização baseados em analogias com os processos de seleção natural e genética evolucionária [37]. A essência do método consiste na evolução de uma população de indivíduos (cromossomos), os quais representam possíveis soluções para um problema específico. A melhor solução é atingida através de um processo de seleção competitivo, envolvendo cruzamentos e mutações [38]. Desde a formulação inicial de Holland, em 1975, os AGs vêm sendo empregados para solucionar problemas complexos, envolvendo funções objetivo mal comportadas com descontinuidades, não-diferenciáveis e multimodais (muitos máximos locais) [37][39].

Em um AG, os cromossomos são representados por códigos, compondo um alfabeto ou *string*, geralmente formado por dígitos binários (0 e 1). A codificação binária é considerada a codificação clássica em AGs e é a mais comumente utilizada, mas é possível encontrar outros

alfabetos, como os de ponto flutuante (codificação real), que vem sendo muito utilizada por suas vantagens em relação a codificação binária [39][40].

Os cromossomos são avaliados por uma função de avaliação para determinar a sua aptidão. Dessa avaliação resulta um valor, chamado de grau de aptidão, ou simplesmente aptidão. Quanto maior a aptidão melhor a qualidade da solução encontrada. Essa informação é utilizada para gerar populações mais aptas, através do processo de seleção, ao longo das gerações do AG. O processo de geração de novas populações é repetido iterativamente até que seja alcançado um critério de parada, como por exemplo: atingir um dado número de gerações ou avaliações, perda de diversidade (população com indivíduos muito semelhantes) e convergência do processo de otimização (não ocorre melhora significativa na solução durante um dado número de gerações).

A representação do AG na forma de fluxograma está ilustrada na Figura 3.1.

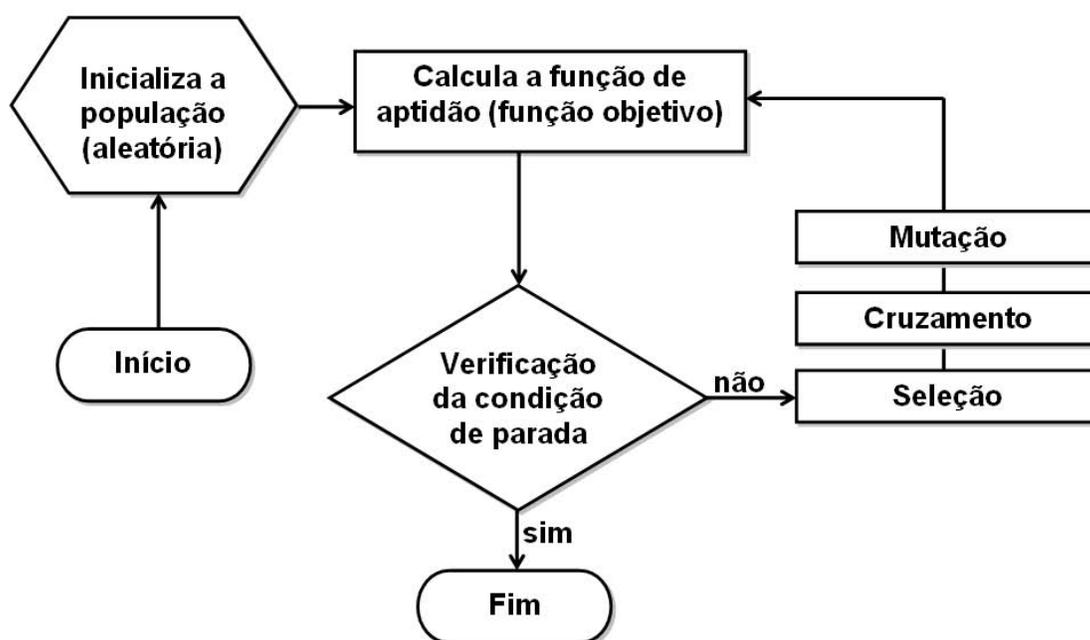


Figura 3.1: Diagrama do algoritmo genético.

3.1.1.1 Tipos de Codificação para os Algoritmos Genéticos

Cada indivíduo de uma população é uma representação codificada de uma possível solução do problema que se deseja resolver. As representações mais utilizadas dos indivíduos do AG são as codificações binária e real. No primeiro caso, são utilizados conjuntos de valores binários (bits), 0's e 1's, para representar os indivíduos do AG. A codificação binária é ideal para problemas nos quais as variáveis são discretas. Se as variáveis são contínuas, uma

conversão será necessária [41]. O processo de conversão pode ser complexo, dependendo do número de bits usados para representar cada parâmetro do problema que está sendo otimizado. Um grande número de bits aumenta a precisão das respostas obtidas pelo AG, porém torna mais complexo o processo de conversão e assim, mais lento o seu processamento. Por outro lado, a codificação real utiliza diretamente valores reais para representar os parâmetros do AG, sendo desnecessário o processo de conversão. As vantagens da codificação real em relação à codificação binária são destacadas a seguir [38]:

- a possibilidade de trabalhar com grandes domínios para as variáveis, o que é difícil atingir em implementações binárias, onde o aumento do domínio pode significar sacrifício da precisão;
- a capacidade de explorar a “gradualidade” das funções com variáveis contínuas. O conceito de gradualidade refere-se ao fato de que pequenas alterações nas variáveis correspondem a pequenas mudanças nas funções. É interessante explorar gradualmente o domínio do problema, pois a solução obtida pode ser melhor através de uma pequena mudança nos parâmetros. A codificação binária não permite essa opção de evoluir gradualmente;
- a capacidade de ajuste local das soluções, como consequência do item anterior;
- a representação das soluções é muito próxima da formulação natural de muitos problemas, ou seja, não há diferenças entre o genótipo (codificação) e o fenótipo (espaço de busca); e
- os processos de codificação/decodificação, necessários nos alfabetos binários, são evitados, aumentando a velocidade dos AG’s.

Cada bit na codificação binária e cada número real na codificação real são denominados de gene. Neste trabalho foi utilizada a codificação real para o problema de identificação proposto, sendo a partir de agora tratada apenas desse tipo de codificação.

3.1.2 Operadores Genéticos

3.1.2.1 Seleção

Seleção é o processo no qual os indivíduos com melhor valor de aptidão têm a maior probabilidade de gerar um ou mais descendentes para a geração seguinte. O objetivo principal do operador de seleção é copiar soluções mais adequadas, eliminando-se soluções de baixa

aptidão. Geralmente nesse processo, o tamanho da população é mantido constante [42]. A seleção é realizada através dos seguintes passos:

1. Identificar as soluções mais aptas da população;
2. Realizar múltiplas cópias das soluções mais aptas;
3. Eliminar soluções de baixa aptidão, permitindo que várias cópias das soluções mais aptas possam ser inseridas na população.

As melhores soluções são armazenadas em uma lista de soluções escolhidas ou lista de soluções, que será utilizada para realizar as operações de cruzamento e mutação.

O operador de seleção é um componente essencial de algoritmos genéticos. A literatura identifica pelo menos cinco principais mecanismos de seleção: proporcional (roleta), por torneios, com truncamento, por normalização linear e por normalização exponencial [43]. Um mecanismo de seleção é caracterizado pela pressão seletiva ou intensidade de seleção que o mesmo introduz no algoritmo genético. A definição de intensidade de seleção empregada em genética é a variação na aptidão média da população induzida pelo método de seleção [43]. A expressão da intensidade de seleção I é dada por:

$$I = \frac{M^* - M}{\sigma} \quad (3.1)$$

onde M é a aptidão média da população, M^* é o valor esperado da aptidão média após a seleção, e σ é o desvio padrão dos valores de aptidão da população antes da seleção.

A seleção proporcional é tipicamente implementada por uma roleta onde cada indivíduo é representado por uma fatia proporcional a sua aptidão relativa. A probabilidade de um indivíduo ser selecionado é simplesmente proporcional ao seu valor de aptidão, isto é:

$$p_i = \frac{f_i}{NM} \quad (3.2)$$

onde p_i é a probabilidade de seleção de um indivíduo i , f_i é a aptidão do mesmo e N é o tamanho da população. Demonstra-se que a intensidade de seleção é dada por [43]:

$$I = \frac{\sigma}{M}, \quad (3.3)$$

isto é, a pressão seletiva é dada pela razão entre o desvio padrão das aptidões e a média das mesmas. O método de seleção proporcional apresenta dois problemas: existência de super-indivíduos e competição próxima. O primeiro ocorre quando um indivíduo apresenta uma

aptidão bem maior que a dos restantes, o que determinará uma convergência prematura do algoritmo. O segundo problema ocorre quando indivíduos apresentam aptidões semelhantes, mas não idênticas; nesse caso, a intensidade de seleção pode ser bem menor do que a desejável.

No método de seleção por torneios, um grupo de t indivíduos é aleatoriamente escolhido, e o indivíduo de melhor aptidão é selecionado. A intensidade de seleção é dada neste caso pela solução da seguinte equação integral [43]:

$$I = \int_{-\infty}^{\infty} t \cdot x \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left(\int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \right)^{t-1} dx \quad (3.4)$$

As variáveis de integração x e y representam os valores de aptidão da população. Assume-se uma distribuição Gaussiana de valores de aptidão por indivíduos. A partir da solução numérica da equação acima, observa-se que a pressão seletiva aumenta a medida em que o número de indivíduos envolvidos no torneio, t , aumenta.

No mecanismo de seleção por truncamento, dado um limiar T , apenas os T melhores indivíduos podem ser selecionados. Cada um desses indivíduos apresenta a mesma probabilidade de seleção. Demonstra-se que a intensidade de seleção é dada por [43] :

$$I = \frac{1}{T} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{f_c^2}{2}} \quad (3.5)$$

onde f_c é o valor da menor aptidão entre os T melhores indivíduos. Observa-se, para esse método de seleção, que a pressão seletiva diminui a medida em que T aumenta.

No método de seleção por normalização linear, os indivíduos são inicialmente ordenados de acordo com sua aptidão. A seguir, esses valores de aptidão são alterados de acordo com a posição relativa de cada indivíduo, dada pela equação:

$$A_i = \min + \frac{\max - \min}{n - 1} \times (i - 1) \quad (3.6)$$

onde \max é o valor de aptidão assinalado ao melhor indivíduo e \min é o valor de aptidão assinalado ao pior indivíduo. Esses dois valores são determinados pelo usuário, mas a forma original deste método prevê que as condições $\max = 2 - \min$ e $\min \geq 0$ devam ser atendidas [43]. Os demais indivíduos têm valores de aptidão linearmente distribuídos entre \min e \max , de acordo com sua posição relativa na ordenação ($i = 1$ corresponde ao pior elemento). A intensidade de seleção é dada por:

$$I = (1 - \min) \frac{1}{\sqrt{\pi}} \quad (3.7)$$

A equação acima mostra que a pressão seletiva diminui a medida em que min aumenta.

O método de seleção por normalização exponencial diferencia-se da normalização linear pelo fato das probabilidades de seleção de cada indivíduo seguirem uma função exponencial. Essa probabilidade é dada por:

$$p_i = \frac{c - 1}{c^N - 1} c^{N-i}; \quad i \in \{1, \dots, N\} \quad (3.8)$$

onde c determina o grau de ‘exponencialidade’ da função, podendo variar de 0 até 1. Quanto mais próximo de 1, menor a ‘exponencialidade’. A intensidade de seleção é dada por:

$$I = \frac{\ln(k)}{-2,548 - 1,086\sqrt{k} + 0,4028 \ln(k)} \quad (3.9)$$

onde $k = c^N$. Através da característica de I em função de c dada pela equação acima, verifica-se que a pressão seletiva diminui a medida em que c aumenta [43].

Neste trabalho, a estratégia de seleção utilizada foi a seleção por torneio, por ser simples, fácil de configurar e apresentar bons resultados, por manter a varibilidade da população e assim evitar convergência prematura do AG.

3.1.2.2 Cruzamento

Indivíduos selecionados (e reproduzidos na população seguinte) são recombinados através do operador de cruzamento (com uma probabilidade p_c). O operador de cruzamento é considerado a característica fundamental dos AGs. Pares de genitores são escolhidos aleatoriamente da população, baseado na aptidão, e novos indivíduos são criados a partir da troca do material genético. Os descendentes serão diferentes de seus pais, mas com características genéticas de ambos os genitores.

O tipo de cruzamento menos complexo dentro da codificação real é o cruzamento simples, onde são escolhidos aleatoriamente dois indivíduos de uma geração e esses geram os dois novos cromossomos. Funciona como o cruzamento de um ponto bastante comum em codificação binária, onde o ponto de cruzamento é escolhido aleatoriamente (desde que não sejam os pontos extremos do cromossomo). Cada novo cromossomo possuirá metade dos genes de cada pai. Por exemplo, têm-se dois cromossomos pais cada um com 10 genes diferentes. Então é obtido um valor aleatório da posição de cruzamento. Se o ponto de corte for a posição 5 nos cromossomos, então, o primeiro filho terá os genes 1, 2, 3, 4 e 5 do primeiro pai e os genes 6, 7, 8, 9 e 10 do segundo pai. Já o segundo filho terá os genes de 1 a 5 provenientes do segundo pai e o restante dos genes do primeiro pai [39][44].

Dois outros métodos de cruzamento que são bastante simples são aqueles que utilizam a média simples e a média geométrica de dois cromossomos pais para gerar um novo cromossomo. Abaixo as fórmulas do cruzamento por média simples (3.10) e por média geométrica (3.11):

$$c = \frac{(p_1 + p_2)}{2} \quad (3.10)$$

$$c = \sqrt{p_1^2 * p_2^2} \quad (3.11)$$

O cruzamento aritmético é realizado através da aplicação de duas fórmulas [39][44]:

$$c_1 = \beta * p_1 + (1 - \beta) * p_2 \quad e \quad c_2 = (1 - \beta) * p_1 + \beta * p_2 \quad (3.12)$$

onde c_1 e c_2 são os cromossomos filhos, p_1 e p_2 são os cromossomos pais e $\beta \in U(0, 1)$, sendo $U(0, 1)$ uma distribuição uniforme.

No cruzamento heurístico é necessário ter conhecimento do valor da função aptidão dos pais. Este método consiste em gerar um cromossomo filho partir de uma interpolação linear entre os pais usando a informação da aptidão. O cruzamento heurístico favorece o pai com maior aptidão. Dados dois cromossomos pais p_1 e p_2 , onde o cromossomo p_1 possui o valor de aptidão superior ao p_2 , o novo cromossomo é gerado a partir da seguinte fórmula [39][44]:

$$c = p_1 + r(p_1 - p_2) \quad (3.13)$$

onde o valor r varia de 0 a 1 e a aptidão do indivíduo p_1 é maior que a aptidão do indivíduo p_2 ($f(p_1) > f(p_2)$).

Pode ser citado também o cruzamento BLX- α , que consiste em gerar um novo cromossomo a partir da Equação (3.14) [45]:

$$c = p_1 + \beta(p_2 - p_1) \quad (3.14)$$

onde c é o cromossomo gerado, p_1 e p_2 são os cromossomos pais e $\beta \in U(-\alpha, 1 + \alpha)$. α é um pequeno valor que estende os limites para a definição de c . A Equação (3.14) é aplicada a cada par de genes de p_1 e p_2 .

Além dos modelos de cruzamento para codificação real citados acima existem outros apresentados na literatura como o cruzamento BGA linear, o cruzamento FCB (fuzzy connectives based) entre outros [46].

Neste trabalho, o método de cruzamento aritmético foi o escolhido para implementação do AG. Esse método é de fácil implementação e é eficaz na pesquisa do espaço de busca. Com

esse método, o AG mantém a regularidade na obtenção de soluções mais aptas varrendo o espaço de busca de forma contínua entre os genes de dois cromossomos selecionados.

3.1.2.3 Mutação

Mutação é um operador exploratório que tem por objetivo aumentar a diversidade na população. Esse operador impede que o algoritmo tenha convergência muito rápida, procurando evitar sua estabilização em regiões de mínimos locais [37]. A probabilidade de mutação p_m controla o processo de mutação, sendo aplicado a cada gene de um cromossomo, determinando se um gene será modificado. Existem diversas formas de se realizar a mutação quando a codificação real está sendo utilizada, tais como mutação uniforme, mutação gaussiana, mutação *creep*, mutação não-uniforme e não-uniforme múltipla [39][44].

Em [38] é proposta a mutação uniforme que é bastante simples. A mutação uniforme é efetuada através da substituição do gene selecionado do cromossomo por outro gene gerado aleatoriamente, segundo uma distribuição uniforme, entre os limites mínimo e máximo permitidos [39][44].

Já a mutação gaussiana é realizada substituindo-se o gene selecionado por outro gerado a partir de uma distribuição normal $N(x, \sigma^2)$ de média igual ao valor de gene x a ser substituído e de variância σ^2 . O valor de σ geralmente apresenta valores acima de 0,5 e alguns trabalhos propõem a sua redução ao longo das gerações [39][44][47].

Na mutação *creep*, acrescenta-se ou subtrai-se do gene um pequeno número aleatório obtido de uma distribuição $N(0, \sigma^2)$ onde a variância assume um valor pequeno [44].

A mutação não-uniforme consiste na simples substituição de um gene por um número extraído de uma distribuição não-uniforme, onde os valores possuem uma probabilidade diferente de ocorrerem. A mutação não-uniforme múltipla consiste em aplicar a mutação não-uniforme em todos os genes do cromossomo selecionado [39][44].

A mutação Discreta modal é apresentada pelas expressões [47, 48]:

$$\gamma = \sum_{K=0}^{\pi} \alpha_K B_M^K \quad (3.15)$$

onde,

$$\pi = \left\lfloor \frac{\log(\text{rang}_{min})}{\log(B_m)} \right\rfloor, \quad (3.16)$$

$B_m > 1$ é um parâmetro conhecido como base da mutação, rang_{min} é o menor limite da faixa de mutação relativa e $\lfloor \cdot \rfloor$ é a operação *floor*.

No AG implementado nesse trabalho foi utilizada a mutação gaussiana, que é simples de configurar além de apresentar um bom comportamento na exploração do espaço de busca.

3.1.2.4 Elitismo

O operador de elitismo mantém as melhores soluções encontradas previamente para as gerações seguintes, garantindo que a melhor solução não se deteriore nas gerações seguintes.

Uma forma de implementar elitismo é copiar diretamente as $n\%$ soluções da população atual para população seguinte. O resto das $(100 - n)\%$ das soluções é gerada usando os operadores genéticos usuais sobre a população atual. Dessa forma, as melhores soluções passam diretamente para a próxima população, além de participar da criação do resto dos indivíduos da população.

Outra forma de elitismo consiste em criar a população seguinte a partir da população atual usando os operadores genéticos usuais, e escolher as melhores N soluções de ambas populações.

Vale destacar que caso sejam utilizados muitos elementos dentro do elitismo, o algoritmo pode convergir precocemente e assim podendo ficar preso a um máximo local.

No AG implementado nesse trabalho, o elitismo faz parte do mecanismo da técnica de otimização multiobjetivo utilizada e é semelhante à segunda técnica de elitismo descrita acima, na qual são usadas duas populações. Para otimização multiobjetivo com AG, demonstra-se que o elitismo aumenta o desempenho do processo de otimização, conforme descrito na Seção 3.3.

3.1.3 Parâmetros genéticos

O comportamento dos Algoritmos Genéticos é influenciado pelos seguintes parâmetros:

1. Tamanho da População: O tamanho da população estabelece o número de cromossomos na população. Esse parâmetro afeta diretamente o desempenho e a eficiência dos AGs. Trabalhando com uma população de poucos indivíduos o desempenho pode ser comprometido visto que o espaço de busca coberto será pequeno. Já uma população grande oferece uma cobertura representativa do domínio do problema e ainda evita o problema de uma convergência prematura para soluções locais. Porém, trabalhar com uma população de tamanho considerável acarreta em uma necessidade maior de recursos computacionais e também de um consumo de tempo maior;
2. Probabilidade de Cruzamento: A rapidez com que novas estruturas são introduzidas na

população depende da probabilidade de cruzamento. Quanto maior for essa probabilidade, mais rapidamente novas estruturas serão introduzidas. Porém, se esse parâmetro for muito alto, a maior parte da população será substituída, e pode ocorrer perda de estruturas de alta aptidão. Com um valor baixo, o algoritmo pode tornar-se muito lento. Geralmente, a probabilidade de cruzamento é configurada antes da execução do AG e mantida constante ao longo das gerações. Mas o valor ótimo desse parâmetro pode depender do problema que está sendo otimizado e da fase de execução do AG. Assim, é interessante ter uma forma de ajuste automático desse parâmetro, associada a evolução das soluções do AG. Para este trabalho foi desenvolvido um método de ajuste desse parâmetro, assim como para a probabilidade de mutação, que será descrito no Capítulo 4;

3. Probabilidade de Mutação: é responsável por determinar a probabilidade em que uma mutação ocorrerá. A probabilidade de mutação evita que a população entre em estagnação, possuindo sempre a mesma informação genética. Com uma probabilidade muito alta, a busca tende a se tornar estritamente aleatória, além de aumentar a chance de que uma boa solução do problema seja destruída. A melhor taxa de mutação varia de acordo com o problema em questão;
4. Intervalo da Geração: Determina a percentagem da população que será substituída na geração seguinte. Caso o valor do intervalo de geração seja muito alto, grande parte da população poderá ser substituída, aumentando assim as chances de se perder indivíduos com alta aptidão. Porém, caso o valor seja muito baixo, o algoritmo pode se tornar muito lento.

3.2 Otimização Multiobjetivo

Muitos projetos de engenharia ou problemas de tomada de decisão envolvem otimização simultânea de múltiplos objetivos. O princípio de otimização multiobjetivo é diferente do processo de otimização mono-objetivo (que possui apenas um único objetivo). Em uma otimização mono-objetivo, pretende-se encontrar a melhor solução de projeto, que corresponde ao valor máximo ou mínimo da função objetivo. Por outro lado, na otimização multiobjetivo com objetivos conflitantes não existe uma única solução ótima. A interação entre diferentes objetivos origina um conjunto de soluções relacionadas, largamente conhecidas como soluções ótimas de Pareto. Já que nenhuma dessas soluções de Pareto pode ser identificada como a melhor solução dentre todas as soluções de Pareto sem uma consideração adicional, o objetivo na otimização multiobjetivo é encontrar quantas soluções ótimas de Pareto for possível. Uma

vez essas soluções sejam encontradas, geralmente são usados processos de tomada de decisão auxiliares para escolher a solução final do processo de otimização a partir dessas soluções.

3.2.1 Problema de Otimização Multiobjetivo

Um problema de otimização multiobjetivo (MOOP, do inglês *Multi-Objective Optimization Problem*) possui várias funções objetivo a serem maximizadas ou minimizadas. Além disso, possui restrições que devem de ser satisfeitas por qualquer solução factível. O enunciado geral para um problema de otimização multiobjetivo é [42]:

$$\left. \begin{array}{ll} \text{Minimizar/Maximizar} & f_m(\mathbf{x}), \quad m = 1, 2, \dots, M; \\ \text{Sujeita às restrições} & g_j(\mathbf{x}) \geq 0, \quad j = 1, 2, \dots, J; \\ & h_k(\mathbf{x}) = 0, \quad k = 1, 2, \dots, K; \\ & x_i^{(L)} \leq x_i \leq x_i^{(U)}, \quad i = 1, 2, \dots, n; \end{array} \right\} \quad (3.17)$$

onde \mathbf{x} é o vetor de n variáveis de decisão $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$. Os valores $x_i^{(L)}$ e $x_i^{(U)}$ representam os valores mínimos e máximos respectivamente da variável x_i . Esses limites definem o espaço de variáveis de decisão ou espaço de decisão D . Ao longo deste trabalho, serão utilizados os termos ponto e solução para designar um vetor solução \mathbf{x} .

As J desigualdades (g_j) e as K igualdades (h_k) são chamadas de funções de restrição. Uma solução \mathbf{x} factível será aquela que satisfaça as $J+K$ funções de restrição e os $2n$ limites. Caso contrário a solução será não factível. O conjunto de todas as soluções factíveis forma a região factível ou espaço de busca S .

Cada uma das M funções objetivo $f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_M(\mathbf{x})$ pode ser individualmente maximizada ou minimizada. Porém, para trabalhar com os algoritmos de otimização, é necessário converter todas as funções objetivo para serem maximizadas ou minimizadas. Em otimização multiobjetivo, o vetor de funções objetivo $\mathbf{f}(\mathbf{x})$ conforma um espaço multidimensional chamado espaço de objetivos Z , além do usual espaço de variáveis de decisão. Essa é uma diferença fundamental em relação à otimização mono-objetivo, cujo espaço de objetivos é unidimensional. Para cada solução \mathbf{x} no espaço de decisão, existe no espaço de objetivos uma função $\mathbf{f}(\mathbf{x})$, como mostrado na Figura 3.2. O mapeamento acontece então entre um vetor \mathbf{x} (n -dimensional) e um vetor $\mathbf{f}(\mathbf{x})$ (M -dimensional). Por exemplo, se cada elemento de \mathbf{x} e $\mathbf{f}(\mathbf{x})$ são números reais, então $\mathbf{f}(\mathbf{x})$ estaria mapeada como $\mathbf{f}(\mathbf{x}) : \mathbb{R}^n \rightarrow \mathbb{R}^M$.

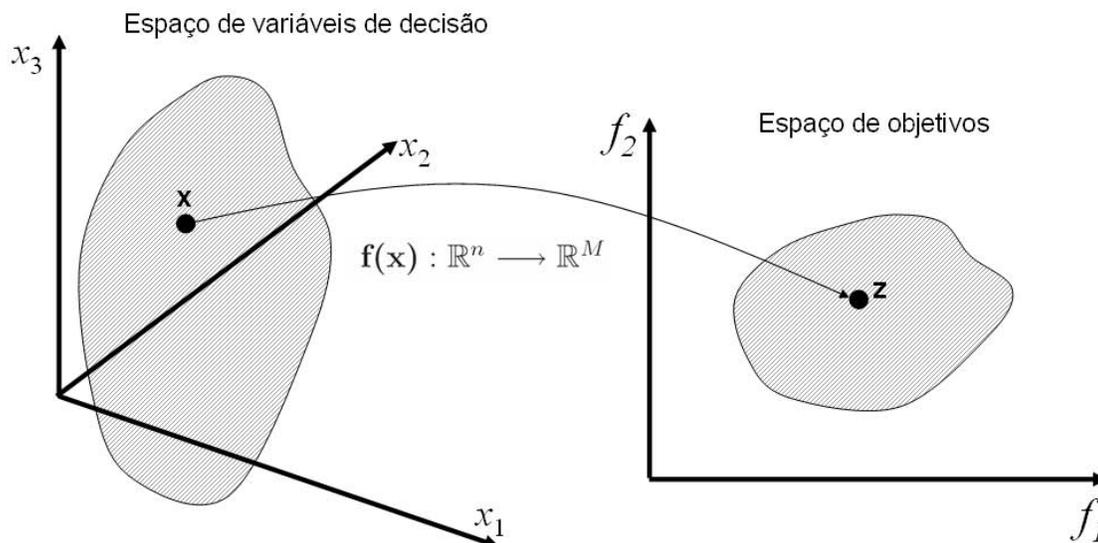


Figura 3.2: Representação do espaço de variáveis de decisão e o correspondente espaço de objetivos.

3.2.2 Soluções Ótimas de Pareto

A otimização multiobjetivo será ilustrada por um exemplo mostrado na Figura 3.3. Nessa figura são mostrados dois objetivos - taxa de acidente e o custo - os quais devem ser minimizados. O ponto A representa uma solução que apresenta custo mínimo, mas é a solução de mais alta propensão à acidentes. Por outro lado, o ponto B representa uma solução de alto custo, mas é a solução de menor propensão à acidentes. Se ambos os objetivos (custo e taxa de acidentes) são objetivos importantes de projeto, não é possível dizer realmente se a solução A é melhor que a solução B, ou vice-versa. Essas soluções são melhores que outras em um objetivo, mas são piores em outro objetivo em relação às mesmas soluções. De fato, existem outras soluções, como a solução D, que possuem valores intermediários de taxa de acidente e custo, sendo de menor custo que a solução B apesar de apresentar maior taxa de acidente que esta mesma solução, e menor taxa de acidentes que a solução A, no entanto maior custo que essa mesma solução. Assim, não se pode concluir sobre uma hierarquia absoluta de soluções A, B, D ou, qualquer outra solução ao longo da linha pontilhada que liga esses pontos, sem alguma informação adicional. Todas essas soluções são conhecidas como soluções ótimas de Pareto.

Observando a Figura 3.3, pode-se notar que existem soluções não-ótimas de Pareto, como o ponto C. Se for comparada a solução C com todas as outras soluções da frente de Pareto formadas pelas soluções A, B e D, será verificado que a solução D é melhor que a

solução C em ambos os objetivos. A solução C, por esse motivo, é conhecida como solução dominada.

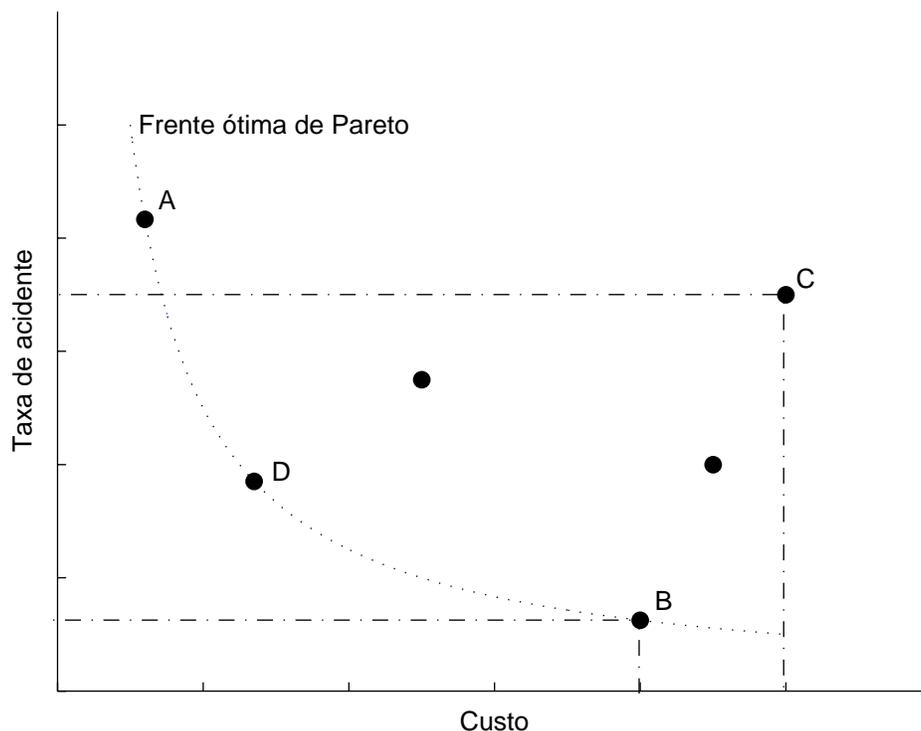


Figura 3.3: Conceito de soluções ótimas de Pareto.

3.2.3 Dominância e Otimalidade de Pareto

O conceito de otimalidade em otimização multiobjetivo trata com um número (ou um conjunto) de soluções, ao invés de uma solução. Baseado nas discussões acima, serão definidas condições para uma solução tornar-se dominada com relação a outra solução e apresentar as condições para que um conjunto de soluções torne-se um conjunto ótimo de Pareto.

Para um problema com M funções objetivo, f_j tal que $j = 1, \dots, M$ e $M > 1$, duas soluções quaisquer $x^{(1)}$ e $x^{(2)}$ (com P variáveis de decisão cada) podem relacionar-se de duas formas: a solução $x^{(1)}$ domina a solução $x^{(2)}$ ou vice-versa, ou não há dominância entre elas. Uma solução $x^{(1)}$ é chamada de dominante em relação a outra solução $x^{(2)}$ se [42]:

1. A solução $x^{(1)}$ é **não pior** que $x^{(2)}$ (considera-se que o operador \prec denota pior e o operador \succ denota melhor), ou $f_j(x^{(1)})$ **não** \prec $f_j(x^{(2)})$ para todos $j = 1, 2, \dots, M$ objetivos;

2. A solução $x^{(1)}$ é **estritamente melhor** que $x^{(2)}$ em ao menos um objetivo, ou $f_j(x^{(1)}) \succ f_j(x^{(2)})$ para ao menos um $j \in \{1, 2, \dots, M\}$

Assim como podem ser definidas soluções locais e globais para otimização mono-objetivo, pode-se definir também conjuntos de soluções ótimas de Pareto locais e globais [42]:

Conjunto ótimo de Pareto local: se para cada elemento x pertencente ao conjunto $\bar{\mathbf{P}}$, não existir uma solução y na sua vizinhança que domine outro elemento do conjunto $\bar{\mathbf{P}}$, então as soluções pertencentes ao conjunto $\bar{\mathbf{P}}$ constituem um conjunto ótimo de Pareto local;

Conjunto ótimo de Pareto global: se não existe solução no espaço de pesquisa que domina qualquer membro no conjunto $\bar{\mathbf{P}}$, então as soluções pertencentes ao conjunto $\bar{\mathbf{P}}$ constituem um conjunto ótimo de Pareto global.

É importante destacar que existe uma diferença entre um conjunto não-dominado e um conjunto ótimo de Pareto. Um conjunto não-dominado é definido no contexto de uma amostra do espaço de pesquisa, isto é, soluções que não são dominadas (de acordo com as definições acima) por qualquer outra solução na amostra do espaço de pesquisa. Um conjunto não-dominado é um conjunto ótimo de Pareto quando a amostra é o espaço de pesquisa inteiro.

3.2.4 Objetivos em Otimização Multiobjetivo

Pode-se dizer que existem primariamente dois objetivos que um algoritmo de otimização multiobjetivo deve tentar alcançar:

1. Guiar a pesquisa para a região ótima de Pareto global;
2. Manter a diversidade da população na frente de Pareto ótima.

A primeira tarefa é um objetivo natural de qualquer algoritmo de otimização. A segunda tarefa é única para otimização multiobjetivo. Já que não existe uma solução no conjunto ótimo de Pareto que pode ser dita melhor que qualquer outra solução, um algoritmo de otimização multiobjetivo tem de encontrar tantas soluções ótimas de Pareto quanto for possível.

3.2.5 Diferença entre Otimização Mono-objetivo e Multiobjetivo

Podem ser citadas pelo menos três importantes diferenças entre otimização multiobjetivo e otimização mono-objetivo [42]:

- Em problemas de otimização com um objetivo, a meta é achar uma solução ótima global (máximo ou mínimo). Em problemas multimodais podem existir mais de um ótimo global. Já em MOOP, achar o conjunto de soluções da Fronteira de Pareto é tão importante quanto preservar a diversidade nesse conjunto. Um algoritmo eficiente para otimização multiobjetivo deve considerar ambos aspectos;
- Um MOOP trabalha com dois espaços (variáveis e objetivos) no lugar de um. Problemas com objetivo simples trabalham unicamente no espaço de variáveis já que procuram apenas uma solução no espaço de objetivos. Novamente, manter a diversidade em ambos espaços complica mais o problema, dado que a proximidade de duas soluções no espaço de variáveis não implica proximidade no espaço de objetivos;
- Os métodos tradicionais de otimização multiobjetivo estão baseados em uma função simples a qual pondera cada objetivo. Podem também tratar cada objetivo separadamente, utilizando os demais objetivos como restrições. Portanto, um MOOP pode ser convertido mediante algumas técnicas, em um problema de otimização simples.

3.2.6 Métodos Clássicos de Otimização Multiobjetivo

Para encontrar as soluções ótimas, os métodos clássicos convertem o MOOP original em um problema equivalente com um único objetivo, reduzindo a otimização vetorial a um problema de otimização escalar. O problema equivalente possui algumas restrições adicionais e para sua solução faz-se necessária a definição de alguns parâmetros, como os valores de pesos das funções objetivos na otimização escalar. Os principais métodos de otimização tradicionais são [40, 49, 50]:

- Somatório de Pesos (Weighted Sum): O Método de Somatório dos Pesos consiste em criar uma função objetivo somando cada objetivo multiplicado por um peso. Esse peso é fornecido como parâmetro. A escolha dos pesos é um problema importante que depende da relevância de cada objetivo. É necessário realizar o escalonamento de cada função objetivo dado que os diferentes objetivos podem ter diferentes magnitudes. Embora esse método seja simples, precisa-se de várias execuções para atingir toda a Fronteira de Pareto. No caso de um MOOP não convexo, este método não é capaz de achar todas as soluções. A aplicação de vetores de pesos uniformemente distribuídos não garante achar um conjunto de soluções uniformemente distribuídas;
- Método de restrições ε (ε -Constraint): Nesse método, o problema multiobjetivo é reformulado considerando qualquer um dos objetivos, e mantendo restritos os demais obje-

tivos com valores definidos pelo decisor, isto é, o método consiste na minimização do objetivo com maior prioridade sujeito à limitação dos demais objetivos. O método de restrições ε pode ser usado para gerar as soluções Pareto ótimas independentemente se o espaço de objetivos for convexo, não convexo ou discreto. Esse método depende da escolha de um vetor de restrições ε que esteja em uma região factível para cada objetivo. Dependendo da escolha desse vetor nenhuma solução poderá se encontrada;

- Programação de Metas (*Goal Programming*): Essa técnica tenta achar soluções que possam atingir uma meta predeterminada para uma ou mais funções objetivo. Caso não exista uma solução factível que alcance as metas para todos os objetivos, esta minimiza os desvios em relação às metas. Para cada objetivo é escolhido um valor meta t pelo usuário. Então o problema é formulado para encontrar uma solução cujo valor em f seja igual a t . Para resolver um problema de programação de metas, cada meta é convertida em uma restrição de igualdade, e procura-se minimizar todos os desvios. Existem várias técnicas para solucionar esses problemas, dentre as quais pode-se citar: programação de metas com Pesos, programação de metas Lexicográficas e programação de metas min-max.

A principal vantagem das técnicas tradicionais é que estas possuem provas de convergência que garantem encontrar as soluções Pareto ótimas. Todas as técnicas descritas neste capítulo fazem a conversão de um MOOP para um problema de objetivo simples. Cada técnica usa uma forma diferente de conversão e introduz parâmetros adicionais. A escolha destes parâmetros afeta diretamente os resultados obtidos. Cada vez que os parâmetros são modificados, tem-se que resolver um novo problema de otimização simples. Portanto, para encontrar N soluções Pareto ótimas, tem-se que solucionar no mínimo N problemas de objetivos simples. Alguns métodos não garantem achar soluções ao longo de toda a Fronteira de Pareto. Se a Fronteira de Pareto não é convexa, o método do Somatório dos Pesos não encontra certas soluções independentemente dos pesos escolhidos. Finalmente, todas as técnicas descritas precisam de parâmetros adicionais tais como pesos, metas, e vetores de restrição. A distribuição uniforme desses parâmetros não garante a diversidade das soluções Pareto ótimas. Os métodos clássicos são sensíveis à convexidade e continuidade da região ótima de Pareto [42].

3.3 Algoritmos Evolutivos para Otimização Multiobjetivo

Outro modo para tratar com problemas de otimização multiobjetivo é utilizar os métodos evolucionários, designados pela sigla MOEA (em inglês, *Multiobjective Evolutionary Algorithms*). Os estudos de algoritmos de pesquisa evolucionários mostraram que esses métodos podem ser usados eficientemente para eliminar a maioria das dificuldades descritas acima, referentes aos métodos clássicos. Já que esses métodos usam uma população de soluções em sua pesquisa, múltiplas soluções ótimas de Pareto podem, em princípio, ser encontradas em uma única execução. Além disso, através de mecanismos para preservar a diversidade da população, os algoritmos de pesquisa evolucionários podem encontrar soluções ótimas de Pareto amplamente diferentes [42].

A primeira implementação de um MOEA foi proposta em [51] por Schaffer. O modelo sugerido foi denominado VEGA (*Vector Evaluated Genetic Algorithm*). Schaffer fez uma modificação nos AGs para avaliar cada objetivo separadamente. Um dos problemas do VEGA, é que esse algoritmo não obtém boa diversidade das soluções na fronteira de Pareto, geralmente convergindo apenas para uma solução ótima. Em 1989, Goldberg criou um procedimento para ordenação de soluções baseado no conceito de dominância [37]. Esse método fornece um valor de aptidão para determinada solução proporcional ao número de soluções que ela domina. Dessa forma, as soluções não dominadas são enfatizadas e terão maior quantidade de cópias na lista de soluções. Com base nas idéias iniciais propostas por Goldberg foram propostas vários modelos de MOEAs. A diferença fundamental dos MOEA em relação aos algoritmos evolutivos tradicionais é o operador de seleção, dado que a comparação entre duas soluções deve realizar-se conforme o conceito de dominância de Pareto. Em alguns métodos, como MOGA (*Multiobjective Genetic Algorithm*) [52] e SPEA (*Strength Pareto Evolutionary Algorithm*) [53], o valor de aptidão é proporcional à dominância da solução. Outros métodos, como NPGA (*Niched-Pareto Genetic Algorithm*) [54], utilizam apenas a dominância de Pareto e não calculam o valor de aptidão.

Os modelos de MOEA podem ser classificados em dois tipos [42]:

1. Não elitista. Esses são modelos, como o nome indica, que não usam nenhum modo de elitismo em suas iterações;
2. Elitista. Modelos que usam algum modo de elitismo. Alguns modelos como SPEA e PESA (*Pareto Envelope-Base Selection Algorithm*) que usam uma população externa na qual são mantidas as soluções não-dominadas encontradas até aquele momento de sua

execução. Outros métodos como NSGA-II (*Non-Dominated Sorting Genetic Algorithm-II*) combinam a população atual com a população posterior para preservar ambas as soluções. O estudo realizado em [55] conclui que o elitismo melhora as soluções encontradas por um modelo MOEA. A partir desse trabalho, os novos algoritmos incorporaram alguma estratégia de elitismo.

Serão apresentadas as características fundamentais apenas das principais pesquisas e/ou implementações no campo dos algoritmos evolutivos multiobjetivo [40, 49].

3.3.1 *Vector Evaluated Genetic Algorithm - VEGA*

Tido como o pioneiro na implementação de algoritmos evolutivos para solução de problemas multiobjetivo, Schaffer desenvolveu o chamado “Vector Evaluated Genetic Algorithms”, mais conhecido como VEGA. Schaffer modificou o software de domínio público GENESIS através da criação de melhorias no procedimento de seleção original, que faz com que o procedimento seja repetido para cada objetivo separadamente, contemplando desta forma a natureza multiobjetivo do problema, até atingir-se um determinado número predefinido de indivíduos para cada objetivo. Em seguida, esses indivíduos são randomicamente sorteados para as etapas de recombinação e mutação. O método foi implementado em combinação com o procedimento de seleção proporcional à aptidão dos indivíduos.

3.3.2 *Multi-objective Optimization Genetic Algorithm - MOGA*

Este algoritmo utiliza um procedimento de ordenamento não-dominado das funções de aptidão. Este procedimento consiste em avaliar a população utilizando os critérios de dominância apresentados, como segue.

1. Os indivíduos não-dominados recebem uma posição ou ordem 1;
2. Os outros indivíduos são posicionados segundo as suas ordens de não-dominância em relação ao restante da população, isto é, para cada indivíduo, o número de soluções que o dominam estritamente é primeiramente determinado na população, logo, a posição no ordenamento desse indivíduo será esse número acrescido da unidade.

Assim sendo, no final deste procedimento de ordenamento poderão existir muitos indivíduos compartilhando a mesma posição no ordenamento. Esse ordenamento é utilizado no procedimento de seleção para escolher ou remover blocos de soluções representados pelos

indivíduos e suas respectivas funções de aptidão, até escolher os indivíduos mais aptos para a reprodução. Esse tipo de atribuição de aptidão por blocos pode acarretar uma grande pressão de seleção, causando convergência prematura do algoritmo.

Para distribuir a população através da região ótima de Pareto e compartilhar os valores da função de aptidão, aplica-se um método de formação de nichos. Essa estratégia mantém a diversidade nos valores da função de aptidão, mas pode não garantir a diversidade no conjunto das variáveis. Assim, o MOGA pode não estar apto a achar as múltiplas soluções em problemas em que diferentes pontos ótimos de Pareto correspondem aos mesmos valores de aptidões. O destaque relevante deste trabalho é o ordenamento dos indivíduos por critérios de dominância.

3.3.3 *Niched-Pareto Genetic Algorithm - NPGA*

O NPGA é um algoritmo genético geracional com sobreposição, isto é, nem todos os indivíduos são substituídos de uma geração para outra. Esse algoritmo utiliza um esquema de seleção baseada na dominância de Pareto. Nesse processo de busca de soluções não-dominadas, dois indivíduos são escolhidos aleatoriamente e comparados com um subconjunto da população (usualmente 10% da população). Quando ambos os competidores são dominados ou não-dominados, um contador de nicho é criado para cada indivíduo na população inteira. O contador é baseado no número de soluções na população com uma certa distância do indivíduo. Assim, a solução que apresentar o menor contador de nicho é selecionada.

3.3.4 *Strength Pareto Evolutionary Algorithm - SPEA*

O SPEA foi proposto por Zitzler e Thiele com o objetivo de introduzir o elitismo nos algoritmos evolutivos multiobjetivos. Inicialmente é gerada uma população aleatória (P_0) de dimensão N_{pop} e um conjunto externo vazio (P_e). As melhores soluções (soluções não-dominadas) de cada geração, denominadas elite, são armazenadas no conjunto P_e . Em cada geração, as soluções não-dominadas da população corrente são copiadas e armazenadas na população externa P_e . Nessa etapa ocorre a união entre as soluções não-dominadas (população de elite) da geração corrente e da geração anterior. Após a união das soluções não-dominadas, essas soluções são comparadas, de forma que as melhores soluções sejam preservadas. À medida que o processo evolui, em alguns problemas, o número de soluções contidas no conjunto de elite torna-se elevado e para restringir o aumento exagerado de soluções não-dominadas no conjunto P_e , determina-se a dimensão máxima desse conjunto.

3.3.5 *Non-Dominated Sorting Genetic Algorithm*

Esse algoritmo foi inicialmente proposto em [56] com o objetivo de trabalhar com os conceitos de dominância e manutenção da diversidade da população simultaneamente. No NSGA utiliza-se um esquema de seleção baseada por ordenamento para privilegiar as soluções não-dominadas correntes, juntamente com um método voltado para a criação de nichos para manter a diversidade da população. A diferença dessa implementação, em relação a um algoritmo genético simples, está apenas no modo com que o operador de seleção é empregado. Tanto o operador de recombinação quanto o operador de mutação são normalmente utilizados nos algoritmos genéticos convencionais. O NSGA difere do algoritmo genético mono-objetivo apenas na maneira como o método de seleção trabalha, uma vez que os operadores cruzamento e mutação permanecem idênticos. Antes do procedimento de seleção ser aplicado, os indivíduos presentes na população corrente são ordenados com base no nível de não-dominância dos indivíduos, isto é, todas as soluções não-dominadas da população corrente recebem valores altos de aptidão. Esta aptidão é a mesma para todos os indivíduos não-dominados, garantindo assim que todos possuam um mesmo potencial reprodutivo. Para manter a diversidade na população as soluções não-dominadas compartilham os seus valores de aptidão segundo suas distâncias Euclidianas (também pode ser adotada a distância de Hamming - número de bits diferentes entre dois indivíduos). Finalmente, divide-se o valor da aptidão de cada indivíduo pelo contador de nichos que é proporcional ao número de vizinhos ao seu redor.

Esse procedimento proporciona a co-existência de pontos ótimos múltiplos na população. O pior valor de aptidão compartilhada na solução da primeira fronteira não-dominada é então guardado para uso posterior. Após a execução do compartilhamento e a modificações das aptidões, os indivíduos não-dominados são ignorados temporariamente para processar o resto dos indivíduos da população. O procedimento para determinar novas soluções não-dominadas (segundo nível) é novamente executado, sendo que agora eles recebem um valor de aptidão um pouco menor que o pior valor de aptidão compartilhada no nível anterior. Uma vez mais o procedimento de compartilhamento é executado entre as soluções não-dominadas do segundo nível e as novas aptidões são calculadas como antes. Este processo é repetido até que todos os membros da população tenham um valor de aptidão compartilhada. A reprodução da população é efetuada utilizando-se a aptidão compartilhada, ou seja, como o primeiro nível de soluções não-dominadas possui as mais altas aptidões, um maior número de cópias dos seus indivíduos será realizado e levará a busca para a fronteira ótima de Pareto. A característica mais importante desse método é que praticamente qualquer número de objetivos pode ser usado para os dois tipos de problemas: maximização ou minimização, bastando mudar o modo como os indivíduos não-dominados são identificados. No entanto, esse método apresentava

alguns inconvenientes como: alta complexidade computacional do processo de ordenamento por não dominância; ausência de elitismo que prejudicava o funcionamento do método; e a necessidade de especificar o parâmetro de compartilhamento σ_{share} , sendo preferível um mecanismo de preservação de diversidade sem a especificação de um parâmetro [1]. Para minimizar esses problemas foi criada uma nova versão do NSGA, chamada de NSGA-II [1]. Essa técnica foi usada neste trabalho na implementação da metodologia de identificação de enlaces DSL e será discutida detalhadamente a sua implementação na próxima seção.

3.3.5.1 *Non-Dominated Sorting Genetic Algorithm - II*

O NSGA-II é um método de otimização multiobjetivo baseado em AG capaz de encontrar soluções bem espalhadas sobre a frente ótima de Pareto e requer baixo esforço computacional. O NSGA-II tem as seguintes características:

1. Usa um mecanismo de elitismo para preservar e usar previamente as melhores soluções encontradas em gerações subsequentes;
2. Usa o conceito de ordenamento por não dominância em AGs;
3. Usa o operador de seleção por torneio de multidão para preservar a diversidade entre as soluções não dominadas nos estágios de execução posteriores, de forma a obter um bom espalhamento das soluções. Nesse mecanismo para preservar a diversidade, se as soluções estão na mesma frente não-dominada, a seleção será baseada na sua distância de multidão, que é uma medida de densidade de soluções na vizinhança de uma determinada solução. A solução situada em uma área menos povoada (com maior distância de multidão) será selecionada.

De acordo de resultados citados na literatura, o NSGA-II tem melhor desempenho que outras técnicas com os mesmos objetivos de preservar o espalhamento das soluções na frente não-dominada [1],[57].

A cada iteração do NSGA-II, uma população pai Φ_n é utilizada para gerar outra população Φ'_n a partir da seleção por torneio e dos operadores genéticos, cruzamento e mutação. Ambas populações tem tamanho N . Para $n = 0$, a seleção de Φ_n é baseada na frente de Pareto que cada solução ocupa ou, de outra forma, no seu nível de não-dominância, sendo 1 o melhor nível (solução na primeira frente de Pareto). Para $n \geq 1$, utiliza-se a seleção por torneio de multidão, que será explicada mais a frente. Em seguida, ambas populações são unidas em uma população $\Phi_n^{\cup} = \Phi_n \cup \Phi'_n$, com tamanho $|\Phi_n^{\cup}| = 2N$. Realiza-se um ordenamento por não dominância sobre Φ_n^{\cup} , obtendo as frentes de Pareto $\mathcal{F}_1, \mathcal{F}_2, \dots$ e todos estes conjuntos são

inseridos na nova população Φ'_n . Dado que apenas N soluções podem ser inseridas, N soluções de Φ'_n são descartadas. Cada conjunto \mathcal{F}_i deve ser inserido na sua totalidade em Φ'_n e isto acontece enquanto $|\Phi'_n| + |\mathcal{F}_i| \leq N$. Ao inserir um \mathcal{F}_j tal que $|\mathcal{F}_j| > N - |\Phi'_n|$, o algoritmo NSGA-II escolhe as soluções de \mathcal{F}_j que estejam melhor espalhadas. Para isso, o algoritmo NSGA-II introduz um método chamado de distância de multidão (*crowding distance*). Uma vez obtidas as distâncias, os conjuntos \mathcal{F}_j são ordenados decrescentemente em relação as suas distâncias, e copia-se as primeiras $N - |\Phi_{n+1}|$ soluções de \mathcal{F}_j para Φ_{n+1} . A Figura 3.4 ilustra uma iteração para o NSGA-II.

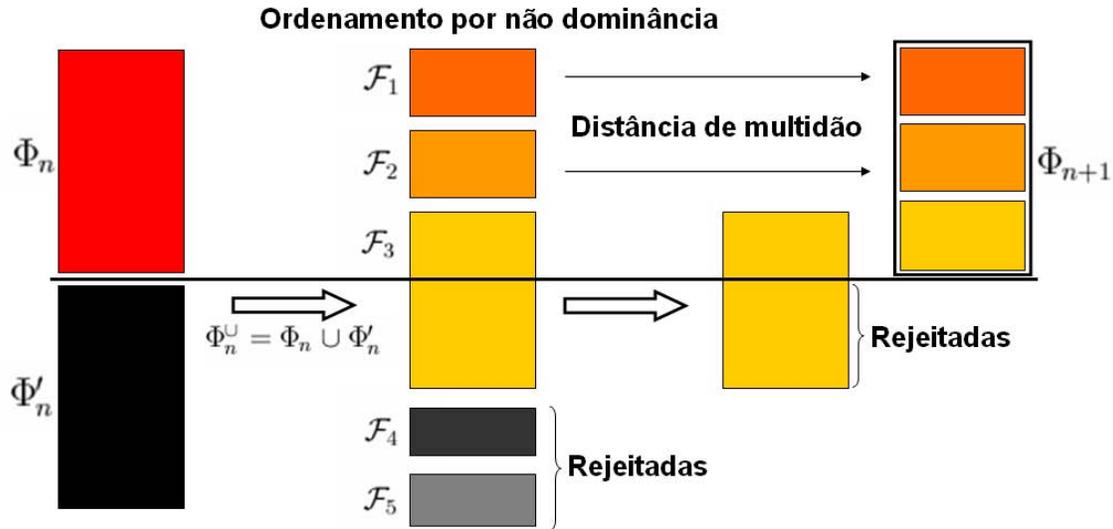


Figura 3.4: Esquema do modelo NSGA-II [1].

3.3.5.2 Distância de Multidão

Para obter uma estimativa da densidade de soluções de um ponto (solução) em particular na população, calcula-se a distância média dos dois pontos de ambos os lados desse ponto envolvendo todos os objetivos. Essa quantidade d_i serve como uma estimativa do tamanho do maior cubóide que inclui o ponto i sem incluir qualquer outro ponto da população, sendo chamada de distância de multidão. Na Figura 3.6, a distância de multidão da i -ésima solução na sua frente Pareto (pontos pretos) é a média do comprimento lateral do cubóide (mostrado em linha tracejada). As soluções extremas em cada objetivo terão um cubóide infinito. O algoritmo descrito na Tabela 3.5 é usado para calcular a distância de multidão de cada ponto na frente \mathcal{F}_j .

I_i^m representa a i -ésima solução na lista ordenada pelo objetivo m ; I_1^m e I_l^m são os elementos da lista com menor e maior valor em um objetivo m ; $f_m^{I_{i+1}^m}$ e $f_m^{I_{i-1}^m}$ são os valores

Figura 3.5: Algoritmo para cálculo da distância de multidão.

Entrada: \mathcal{F}_j (conjunto de soluções na fronteira j), onde $|\mathcal{F}_j| = l$

Saída: d_i para cada solução i em \mathcal{F}_j

1. Para cada solução i em \mathcal{F}_j atribui-se $d_i = 0$;
2. Para cada função objetivo f_m de $m = 1, 2, \dots, M$;
3. Ordenar decrescentemente as soluções por f_m na lista I_m ;
4. Para cada solução extrema (mínimo e máximo) em cada um dos M objetivos;
5. Fazer $d_{I_1^m} = d_{I_l^m} = \infty$;
6. Para as soluções $i = 2, \dots, l - 1$ calcular:
- 7.

$$d_{I_i^m} = d_{I_i^m} + \frac{f_m^{(I_{i+1}^m)} - f_m^{(I_{i-1}^m)}}{f_m^{max} - f_m^{min}} \quad (3.18)$$

dos objetivos das soluções vizinhas à solução i considerando a m -ésima função objetivo. f_m^{max} e f_m^{min} são parâmetros dos limites máximo e mínimo em cada objetivo. A Equação (3.18) assegura que as soluções mais afastadas tenham distâncias de multidão maiores.

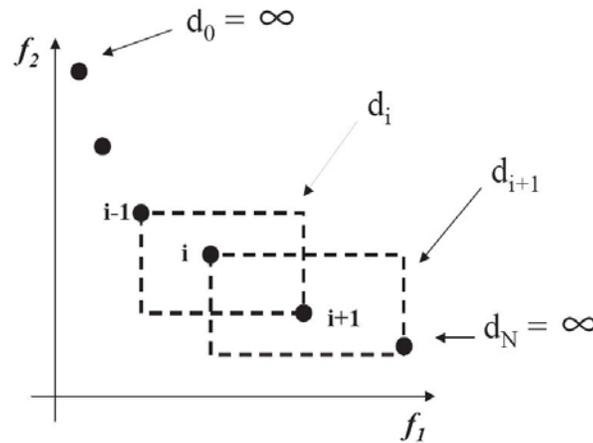


Figura 3.6: Procedimento para cálculo da distância de multidão [1].

3.3.5.3 Operador de comparação de multidão

NSGA-II incorpora uma pequena modificação no método de seleção por torneio, usando o operador comparativo que leva em conta a multidão de uma solução (*crowded tournament selection operator*) (\geq_n). Uma solução i é considerada ganhadora em um torneio contra uma solução j , se:

1. A solução i possui um melhor nível de não dominância;

2. Se ambas soluções estão no mesmo nível, mas i tem uma distância de multidão maior,
 $d_i > d_j$;

Ou em outros termos, o operador (\geq_n) pode ser descrito como:

$$i(\geq_n)j \quad (r_i < r_j \text{ ou } ((r_i < r_j) \text{ e } (d_i > d_j))),$$

onde r_i e r_j representam o índice da frente na qual a solução i e a solução j se encontram, respectivamente. Isto é, entre duas soluções com diferentes níveis de não dominâncias, são escolhidos os pontos com menor nível. Por outro lado, se ambos pontos pertencem a mesma frente, então são escolhidos pontos localizados em uma região com menor número de pontos, que apresenta maior distância de multidão.

3.4 Conclusões do Capítulo

Neste capítulo foram tratados os conceitos principais sobre algoritmos evolucionários, com destaque para os algoritmos genéticos multiobjetivos. Os operadores genéticos selecionados para o AG implementado, foram o cruzamento aritmético, a mutação gaussiana e a seleção por torneio, por serem simples de implementar e apresentarem bom desempenho. Já o AG multiobjetivo escolhido para a metodologia de identificação foi o método NSGA-II, que foi descrito detalhadamente. A partir de testes realizados e do que é descrito na literatura, essa técnica foi escolhida por apresentar melhores resultados que outras técnicas baseadas em elitismo por não dominância e também por ser de fácil implementação. Mas como será visto nos próximos capítulos, apesar da técnica de otimização multiobjetivo escolhida ser de grande importância para a metodologia de identificação, existem outros aspectos que podem ser mais importantes, como as especializações realizadas que se baseiam nas condições específicas da identificação de enlaces DSL.

No próximo capítulo será apresentada a metodologia de identificação de enlaces DSL, que se baseia nas técnicas de análise de linhas transmissão descritas no Capítulo 2 e em processos de otimização, como o AG multiobjetivo descrito neste capítulo e em um processo de otimização local, além de utilizar técnicas analíticas.

Capítulo 4

Identificação de Topologias Usando Computação Evolucionária Multiobjetivo - TIMEC

Neste capítulo e no próximo é apresentada a metodologia proposta para identificação de enlaces DSL, chamada LQ-SUITE. Cada um dos módulos do LQ-SUITE é descrito, começando com TIMEC, apresentado neste capítulo, que utiliza otimização multiobjetivo baseada em algoritmos genéticos e otimização local para identificar a topologia de enlaces. A seguir são descritas as técnicas de codificação/decodificação utilizadas pelo AG e que podem ter grande influência no desempenho da metodologia de identificação. Na seção seguinte, é apresentado o método de ajuste dos parâmetros do AG, que torna a configuração do algoritmo automática e melhora o seu desempenho. Também é apresentada a ferramenta de medição automática, responsável por controlar os equipamentos de medida e fazer o tratamento dos dados obtidos. Finalmente, na última seção são apresentadas as conclusões.

4.1 Ferramenta de Qualificação Identificação de enlaces DSL - LQ-SUITE

A metodologia proposta, LQ-SUITE, usa medidas de SELT e DELT para identificar a topologia de um enlace DSL. Os parâmetros usados pela metodologia são a função de transferência H , o parâmetro de espalhamento S_{11} e a impedância de entrada Z_{IN} do enlace. O objetivo da metodologia proposta é estimar o conjunto de parâmetros de um enlace com n_s

seções, como o enlace mostrado na Figura 4.1, no qual a terceira seção é uma *bridged-tap*. Além de estimar o número de seções n_s do enlace, para cada seção k é necessário estimar seu comprimento l^k , tipo ^{k} (serial ou *bridged-tap*) da seção e bitola (diâmetro do fio) g^k .

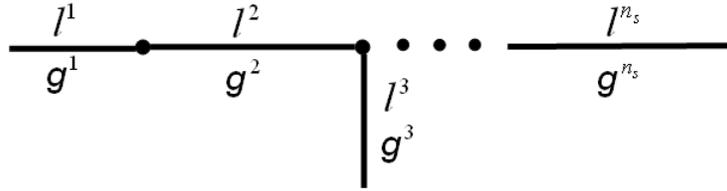


Figura 4.1: Parâmetros do enlace Θ a serem determinados. Nesse exemplo, a terceira seção é uma *bridged-tap* enquanto todas as outras são serials.

O LQ-SUITE é constituído de três módulos, como mostrado na Figura 4.2. O principal deles, o TIMEC (*Topology Identification via Model-based Evolutionary Computation*), utiliza um processo de *analise-por-síntese* baseado em otimização multiobjetivo usando AG e em uma técnica local de otimização. Seu objetivo é obter a topologia completa do enlace a partir das medidas de função de transferência H e do parâmetro de espalhamento S_{11} , com ou sem informação *a priori* do enlace. O segundo módulo é o IRA (*Impulse Response Analysis*) que realiza a análise da resposta impulsiva para obter informação sobre o enlace, que são usadas pelo TIMEC. Essas informações podem ser o comprimento total e o número de *bridged-taps* com seus comprimentos e bitolas. O terceiro módulo é o FSF (*First Section Finder*) que utiliza a curva TDR, obtida a partir da impedância de entrada do enlace, para também estimar informações *a priori* sobre o enlace. Com essa técnica, as informações obtidas podem ser o comprimento total, a topologia da primeira seção do enlace, e o número de *bridged-taps* e suas posições ao longo do enlace. O funcionamento dos módulos IRA e FSF será descrito no Capítulo 5. Enquanto TIMEC é descrito nas próximas seções.

4.2 O método TIMEC

O método TIMEC é baseado em otimização multiobjetivo utilizando algoritmos genéticos multiobjetivos e otimização local. Os objetivos de otimização ou curvas de referência usados na estimação dos parâmetros do enlace são o parâmetro de espalhamento S_{11} e a função de transferência H , obtidos através de medições de SELT e DELT, respectivamente. Eventualmente, a impedância de entrada Z_{IN} pode ser usada no lugar do S_{11} , mas nesse trabalho será assumido que a medição de SELT usada por TIMEC é o parâmetro de espalhamento S_{11} . Os parâmetros de um enlace, como o mostrado na Figura 4.1, serão representados por um

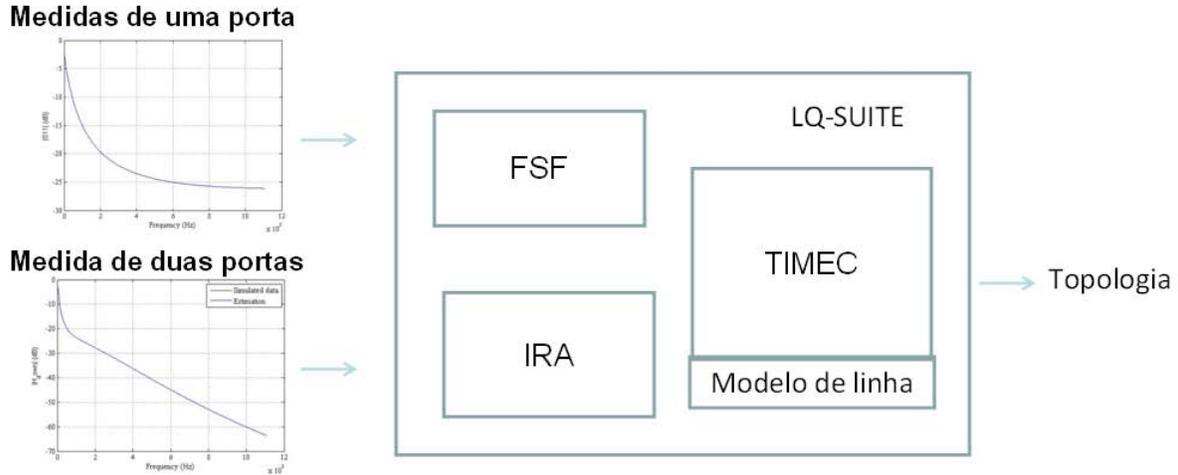


Figura 4.2: LQ-SUITE e seus módulos: TIMEC, IRA e FSF.

vetor Θ dentro do contexto de TIMEC, indicando uma solução do processo de otimização. Para fins de comparação, os parâmetros verdadeiros do enlace serão representados por Θ^\dagger .

Um processo de *análise-por-síntese* é usado de forma que, para uma dada solução candidata Θ , modelos de linha e a teoria ABCD padrão são usados para obter as curvas correspondentes de S_{11} e H , através das Equações (2.10) e (2.11), respectivamente. Esse processo para obtenção de S_{11} e H será representado pelo operador \mathcal{V} . Assim, para uma dada solução Θ :

$$[H, S_{11}] = \mathcal{V}(\Theta) \quad (4.1)$$

Ambos S_{11} e H são comparados às curvas medidas, representadas por \hat{S}_{11} e \hat{H} , usando as funções objetivo, ou funções custos, $V_H(\Theta)$ e $V_{S_{11}}(\Theta)$, respectivamente. Rotinas de otimização pesquisam iterativamente o espaço de soluções para encontrar a melhor solução Θ^* baseada em ambas funções custos. O objetivo final de TIMEC é obter $\Theta^* = \Theta^\dagger$, encontrando assim a topologia correta do enlace. O TIMEC é constituído de dois processos de otimização executados concomitantemente. A otimização multiobjetivo baseada em AG que busca encontrar a solução completa, considerando os diferentes tipos de parâmetros como o número de seções, o tipo, o comprimento e a bitola das seções; e a otimização local que a partir da estrutura topológica (número de seções, tipos e bitolas das seções do enlace) fornecida iterativamente pelo AG, estima o comprimento das seções. Os valores iniciais dos comprimentos do enlace usados pelo otimizador local são obtidos da estimativa fornecida pelo AG. Essas duas partes de TIMEC serão descritas separadamente, mas sempre tendo em mente que elas trabalham

juntas na obtenção da solução ótima.

4.2.1 Otimização Baseada em AG

Diversas rotinas de otimização podem ser aplicadas dentro do contexto de análise-por-síntese proposto, mas AG foi escolhido por ser flexível e bem adaptado para otimização de espaços multidimensionais, com variáveis contínuas e discretas, que apresentam muitos mínimos locais. A utilização de método de pesquisa força bruta para obter a melhor solução Θ^* , testando exaustivamente cada possível solução Θ , não é factível devido a complexidade e grande extensão do espaço de soluções, constituído de variáveis contínuas e discretas. Por outro lado, AG evita a pesquisa força bruta, fazendo uma pesquisa do espaço de soluções mais seletiva, através de uma população de soluções (indivíduos, na terminologia do AG), que tem sua aptidão aumentada iterativamente ao longo das gerações do AG. Por outro lado, a aplicação bem sucedida do AG para um problema depende tipicamente de sua adaptação às especificidades desse mesmo problema. No problema em questão, que trata da identificação de enlaces DSL, o AG otimiza o conjunto de parâmetros Θ , comparando H e S_{11} obtidas a partir de Θ com as curvas de referência \hat{H} e \hat{S}_{11} obtidas de medições, usando as seguintes expressões, respectivamente:

$$V_H(\Theta) = \sum_{k=1}^K \left| H(f_k) - \hat{H}(f_k) \right|^2 \quad (4.2)$$

e

$$V_{S_{11}}(\Theta) = \sum_{k=1}^K \left| S_{11}(f_k) - \hat{S}_{11}(f_k) \right|^2, \quad (4.3)$$

onde f_k é o k -ésimo tom de frequência.

Já que duas funções objetivo são usadas, foi adotado um algoritmo de otimização multiobjetivo, denominado de NSGA-II [1], baseado em seleção por frentes de Pareto, que apresenta melhores resultados que outros algoritmos baseados no mesmo critério de seleção [1] [57]. Ao final do processo de otimização, a melhor solução é selecionada da primeira frente de Pareto \mathcal{F}_1 , usando-se um critério pré-estabelecido. Entre todas as soluções candidatas Θ_r na primeira frente de Pareto, seleciona-se a solução candidata que possui a menor soma dos valores dos objetivos, que pode ser representada pela seguinte operação:

$$\Theta^* = \arg \min_{r \in \mathcal{F}_1} (V_H(\Theta_r) + V_{S_{11}}(\Theta_r)), \quad (4.4)$$

onde r é o índice de uma dada solução na primeira frente de Pareto. Esse critério foi escolhido devido à paridade de relevância de ambos objetivos na identificação de enlaces DSL, sendo

definido, assim, apenas um critério de desempate entre as melhores soluções associadas a cada um dos objetivos utilizados no processo de otimização. Outras alternativas de decisão para definir a melhor solução, considerando a maior relevância de uma função objetivo em particular, podem ser usadas. Mas o critério definido na Equação 4.4 tem sempre levado ao melhor resultado para os testes de TIMEC que serão mostrados no Capítulo 6.

Na Figura 4.3 é mostrado o diagrama de bloco do TIMEC, considerando apenas a parte referente a otimização por AG. As variáveis e constantes mostradas nesse diagrama são descritas na Tabela 4.1.

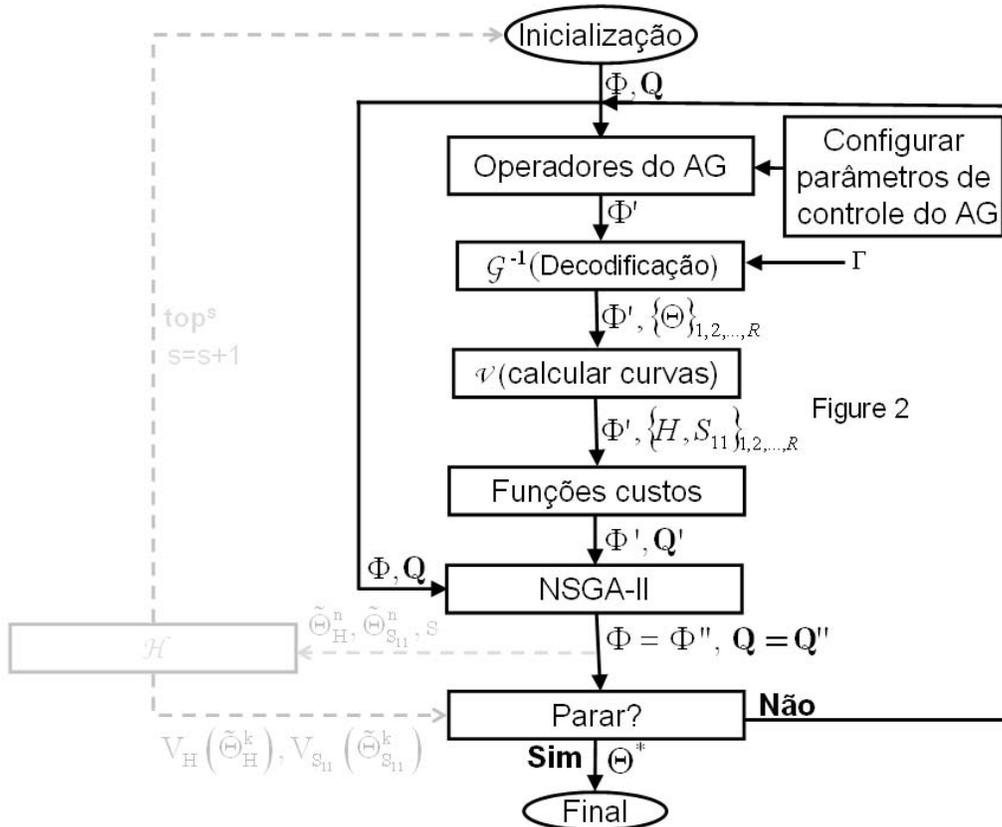


Figura 4.3: Diagrama de bloco do TIMEC com apenas a otimização global.

Na fase inicialização do AG, é gerada uma população inicial Φ , com indivíduos gerados aleatoriamente. O tamanho da população pode ser definido por $R = |\Phi|$. Os valores da função custo para todos os indivíduos de Φ são calculados usando as Equações (4.2) e (4.3) e

Tabela 4.1: Breve descrição do símbolos usados.

Símbolo	Descrição
Parâmetros definidos pelo usuário	
Σ_m	Valor máximo do desvio padrão associado à mutação
P_m	Valor máximo da probabilidade de mutação
P_c	Valor máximo da probabilidade de cruzamento
R	Tamanho das populações do AG (Φ , Φ' e Φ'')
N	Número total de gerações
ΔN	Incremento do número de gerações
δ_s	Limiar de erro usado como condição de parada
N_s	Número máximo de seções
N_b	Número máximo de <i>bridged-taps</i> (pode depender do valor de N_s)
Estruturas de dados e variáveis	
Γ	Armazena informações <i>a priori</i> sobre o enlace (comprimento, bitola e tipo)
p_m	Probabilidade de mutação da geração atual
p_c	Probabilidade de cruzamento da geração atual
σ_m	Desvio padrão para mutação da geração atual
Φ	População do AG, i.e., conjunto codificado de possíveis soluções
Φ', Φ''	Populações temporárias
\mathbf{Q}	Matriz objetivo que armazena os valores das funções custos
$\mathbf{Q}', \mathbf{Q}''$	Matrizes objetivo temporárias
n	Contador para o número de iterações do AG
$\tilde{\Theta}_H$	Indivíduo decodificado, ou solução candidata, que apresenta o menor valor para função custo associada à função de transferência
$\tilde{\Theta}_{S_{11}}$	Indivíduo decodificado, ou solução candidata, que apresenta o menor valor para função custo associada ao parâmetro de espalhamento S_{11}
Θ^*	Topologia final do processo de otimização (a melhor solução encontrada pelo AG)
n_s	Número de seções
n_b	Número de <i>bridged-taps</i>

organizados em uma matriz, chamada de matriz de objetivos:

$$\mathbf{Q} = \begin{bmatrix} V_H(\Theta_1) & V_{S_{11}}(\Theta_1) \\ V_H(\Theta_2) & V_{S_{11}}(\Theta_2) \\ \vdots & \vdots \\ V_H(\Theta_R) & V_{S_{11}}(\Theta_R) \end{bmatrix}$$

com dimensão $R \times 2$.

A Figura 4.3 indica que o NSGA-II melhora iterativamente Φ até produzir a solução final Θ^* . Para cada geração n , os operadores do AG (seleção, mutação e cruzamento) são aplicados à população Φ para gerar outra população Φ' . Os valores da probabilidade cruzamento p_c e da probabilidade de mutação p_m , assim como o desvio padrão σ_m , são atualizados a cada geração através de um algoritmo de adaptação dos parâmetros do AG, que será descrito na Seção 4.2.4. Para calcular as curvas de H e de S_{11} através do operador \mathcal{V} , cada cromossomo deve ser decodificado em seu correspondente vetor de parâmetros Θ . Processos apropriados de codificação e decodificação, representados respectivamente por $\mathcal{G}(\cdot)$ e $\mathcal{G}^{-1}(\cdot)$, são cruciais e serão descritos em detalhes na Seção 4.2.3. Se alguma informação *a priori* sobre o enlace é disponível, então Γ , que armazena informações *a priori* sobre o enlace, recebe essa informação, que é usada na decodificação dos indivíduos do AG. Através do processo de decodificação da população Φ' , $\mathcal{G}^{-1}(\Phi')$, são obtidos os vetores de parâmetros decodificados $\{\Theta'_1, \dots, \Theta'_R\}$, isto é:

$$\mathcal{G}^{-1}(\Phi') = \{\Theta'_1, \dots, \Theta'_R\}.$$

O próximo bloco aplica o operador \mathcal{V} para calcular o parâmetro de espalhamento S_{11} e a função de transferência H correspondentes a cada um dos vetores de parâmetros $\{\Theta'_1, \dots, \Theta'_R\}$, que são comparadas as curvas medidas \hat{H} e \hat{S}_{11} , através das funções custos, obtendo-se assim a matriz \mathbf{Q}' .

Ambas populações, Φ e Φ' , e as suas respectivas matrizes \mathbf{Q} e \mathbf{Q}' , são usadas pelo processo NSGA-II para gerar outra população Φ'' e a sua respectiva matriz objetivo \mathbf{Q}'' . Basicamente, o NSGA-II seleciona os indivíduos considerando as frentes de Pareto que ocupam no agrupamento daquelas populações. A população Φ'' também possui o tamanho igual a R .

Há dois critérios de parada: o número total de gerações a ser executado N , isto é, o AG encerra sua execução sempre que n é igual ao número total de gerações N ; e um critério de parada adicional dado pela expressão $V_H(\tilde{\Theta}) + V_{S_{11}}(\tilde{\Theta}) < \delta_s$, onde δ_s é um valor definido pelo usuário que indica o mínimo valor aceitável da soma das funções custos e $\tilde{\Theta}$ é o enlace decodificado associado a qualquer um dos melhores indivíduos da população do AG na geração atual $n - \mathcal{G}(\tilde{\Theta}_H)$ ou $\mathcal{G}(\tilde{\Theta}_{S_{11}})$. Se pelo menos uma dessas condições de parada não é alcançada,

ambos Φ'' e \mathbf{Q}'' substituem Φ e \mathbf{Q} , respectivamente, como indicado na Figura 4.3, e uma nova geração do AG é processada usando essa nova população Φ . Caso contrário, a melhor solução Θ^* é selecionada da população Φ usando-se a expressão 4.4, como explicado anteriormente.

4.2.2 Otimização local e processo de realimentação de seções

O AG lida bem com problemas que envolvem múltiplas variáveis de diferentes tipos, como no caso tratado neste trabalho, onde tem-se variáveis booleanas, inteiras e reais. Em muitos casos, é rapidamente encontrada a estrutura topológica do enlace, mas os comprimentos são lentamente ajustados até os valores corretos, justamente pelo AG ser um processo de otimização global e o comprimento ser uma variável real. Por outro lado, apesar de um otimizador local ter dificuldade de tratar com a identificação completa do enlace, que envolve aquelas diferentes variáveis, ele é capaz de estimar rapidamente os comprimentos a partir da estrutura topológica correta. Outro aspecto interessante, é a rápida obtenção da estimativa correta da topologia das seções próximas ao CO, que ocorre devido a influência do parâmetro de espalhamento S_{11} obtido a partir dessa porta. No entanto, em alguns casos, as estimativas de seções mais distantes podem ser incorretas, devido também a essa polarização do algoritmo de otimização à porta referente ao CO, sob a influência de S_{11} . Para tirar vantagem da rápida estimativa das seções próximas ao CO e minimizar o erro para seções mais distantes, foi desenvolvido um algoritmo baseado na realimentação de informações sobre o enlace, que será descrito mais a frente. O objetivo de utilizar a técnica de otimização e o algoritmo de realimentação é melhorar o intervalo de confiança, aumentar a precisão e reduzir o tempo de processamento de TIMEC. Essas inclusões ao método TIMEC mostram que a metodologia apresenta grande flexibilidade, podendo ser facilmente estendida para inclusão de melhorias ao processo de identificação de topologias.

Na Figura 4.4 é mostrado o TIMEC em sua configuração completa, que inclui o bloco \mathcal{H} , que representa a otimização local e o processo de realimentação. Esse bloco recebe a melhor solução de S_{11} e a melhor solução de H para otimizá-las em relação a esses respectivos objetivos.

O bloco \mathcal{H} pode ser expandido no diagrama mostrado na Figura 4.5. Especificamente, o melhor de S_{11} , $\tilde{\Theta}_{S_{11}}^n$, e o melhor de H , $\tilde{\Theta}_H^n$, são otimizados separadamente em relação a S_{11} e H , respectivamente, usando o algoritmo de *Levenberg-Marquardt*. Deve-se notar, como explicado anteriormente, particularmente à abordagem desenvolvida, que apenas os comprimentos das seções é que são otimizados, sendo considerados fixos o número de seções, os valores das bitolas e os tipos das seções. Em outras palavras, a técnica global fornece a estrutura topológica do enlace (número de seções e tipos das seções, além das bitolas) à técnica local, que se encarrega

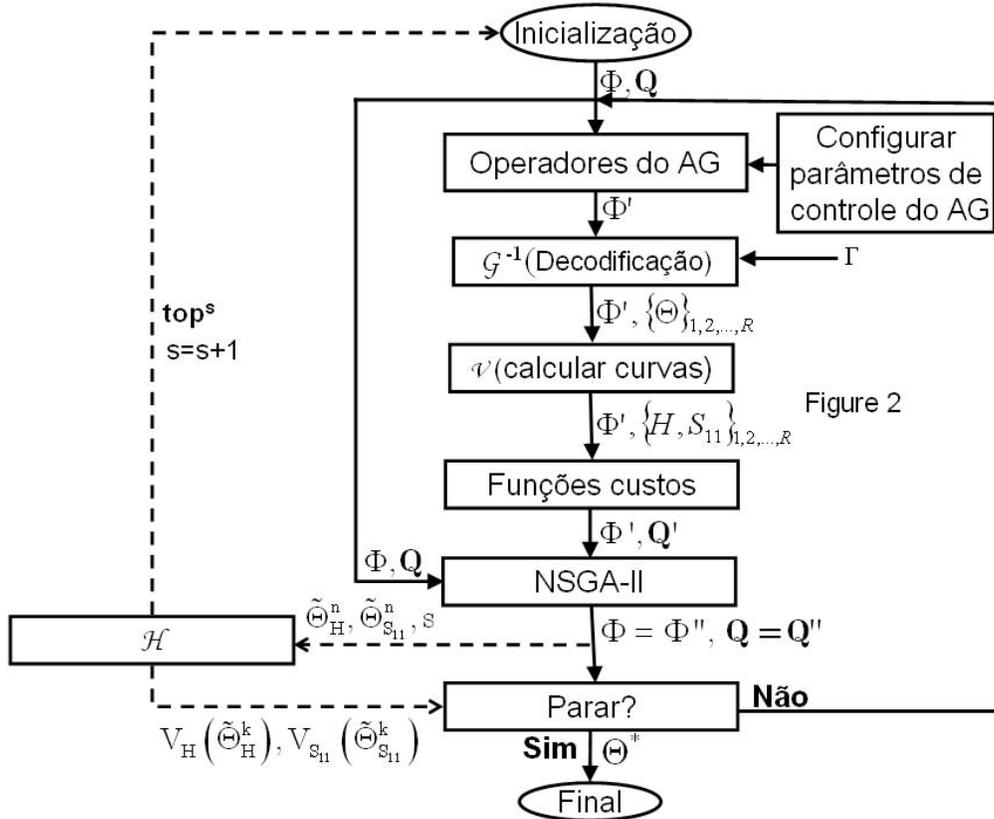


Figura 4.4: Diagrama de bloco do TIMEC com otimização global e local.

de otimizar a única variável contínua da identificação de enlaces (comprimento das seções) que o AG levaria mais tempo em otimizar. Os comprimentos das seções armazenadas nos vetores de parâmetros $\tilde{\Theta}_H^n$ e $\tilde{\Theta}_{S_{11}}^n$ são considerados as estimativas iniciais para a técnica de otimização local. Assim, ao final do processo de otimização local, o resultado será um enlace com a mesma estrutura topológica, mas com comprimentos das seções de valores diferentes dos fornecidos pela técnica global, como é ilustrado na Figura 4.6. Esses dois novos enlaces gerados pela técnica local serão representados pelos vetores de parâmetros $\tilde{\Theta}_H^k$ e $\tilde{\Theta}_{S_{11}}^k$, onde k indica a k -ésima vez que o fluxo de dados do diagrama da Figura 4.4 acionou o bloco \mathcal{H} , considerando que esse é acionado a cada $N_{\mathcal{H}}$ gerações do AG. Assim, a cada intervalo entre $N_{\mathcal{H}}$ gerações (por exemplo, entre as gerações 1 a $N_{\mathcal{H}}-1$), o TIMEC é constituído como na Figura 4.3, de outra forma o TIMEC é como mostrado na Figura 4.4. Os dois novos enlaces gerados $\tilde{\Theta}_H^k$ e $\tilde{\Theta}_{S_{11}}^k$ são comparados aos enlaces obtidos pela última execução do bloco \mathcal{H} , $\tilde{\Theta}_H^{k-1}$ e $\tilde{\Theta}_{S_{11}}^{k-1}$, respectivamente. Se forem soluções melhores, isto é, apresentarem um valor menor para o objetivo que estão relacionadas ($\tilde{\Theta}_H$ para função objetivo de H , por exemplo), substituem as soluções atuais, senão são simplesmente descartadas.

Como pode ser visto na Figura 4.5, além da melhor solução de cada objetivo na geração

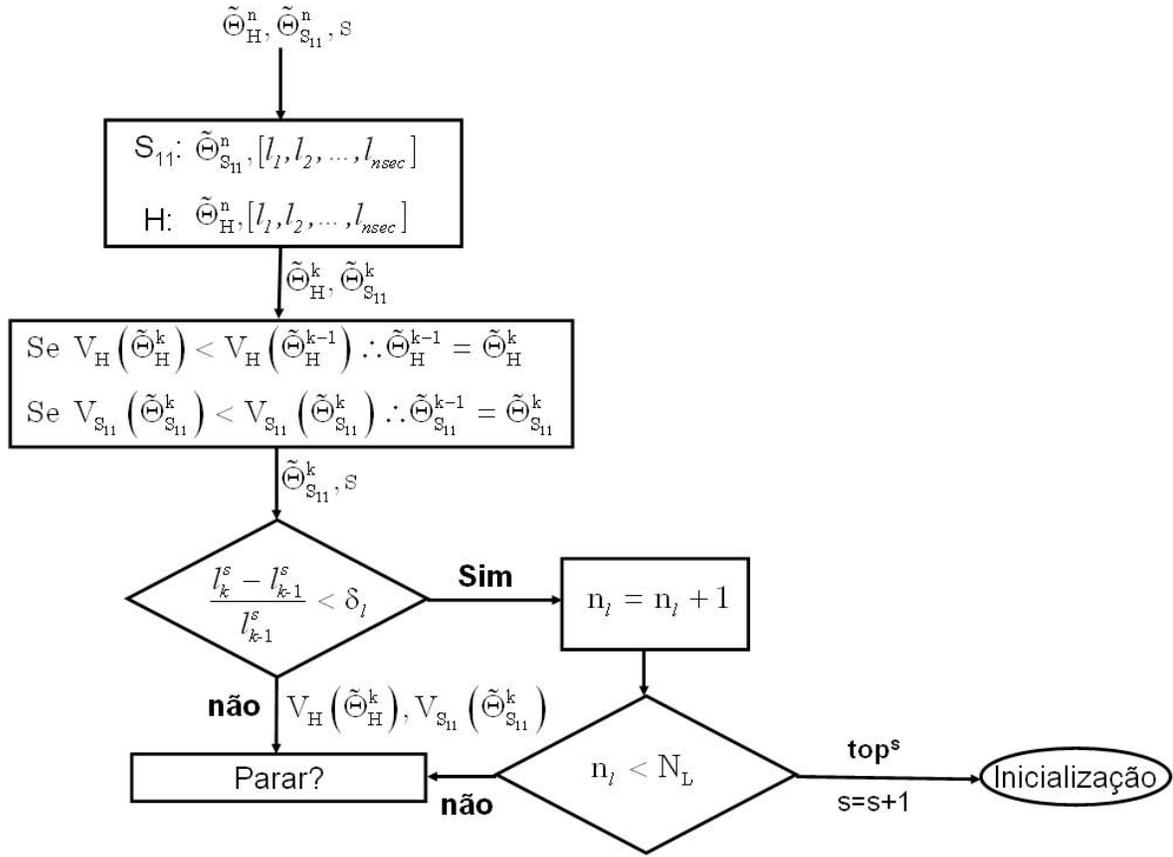


Figura 4.5: Vis\u00e3o expandida do bloco \mathcal{H} de TIMEC.

n do AG, \u00e9 passada tamb\u00e9m um \u00edndice s , que indica a se\u00e7\u00e3o que est\u00e1 sendo testada atualmente pela processo de realimenta\u00e7\u00e3o de se\u00e7\u00f5es. O valor de s tem inicialmente o valor de 1, indicando que a primeira se\u00e7\u00e3o est\u00e1 sendo testada. A cada execu\u00e7\u00e3o k da t\u00e9cnica local, o valor de comprimento da se\u00e7\u00e3o s da melhor solu\u00e7\u00e3o atual de S_{11} \u00e9 comparada com o comprimento da solu\u00e7\u00e3o escolhida na execu\u00e7\u00e3o $k - 1$ da t\u00e9cnica local, como dada na equa\u00e7\u00e3o a seguir:

$$\frac{(l_k^s - l_{k-1}^s)}{l_{k-1}^s} < \delta_l,$$

o valor de δ_l que pode ser configurado pelo usu\u00e1rio e indica se houve varia\u00e7\u00e3o significativa do comprimento da se\u00e7\u00e3o s desde a \u00faltima execu\u00e7\u00e3o do bloco \mathcal{H} . Caso as varia\u00e7\u00f5es sejam menores que o patamar definido, considera-se que o comprimento da s foi mantido constante e um contador n_l \u00e9 incrementado. Caso contr\u00e1rio, o valor do contador n_l \u00e9 levado a 0. Se o valor de n_l alcan\u00e7ar um certo valor N_L , a topologia da se\u00e7\u00e3o s , top^s , \u00e9 considerada encontrada e o todo o processo de otimiza\u00e7\u00e3o do AG \u00e9 inicializado. O valor da topologia \u00e9 armazenado em Γ juntamente com a topologia de outras se\u00e7\u00f5es se houver, isto \u00e9,

$$\Gamma = \Gamma \cup top^s$$

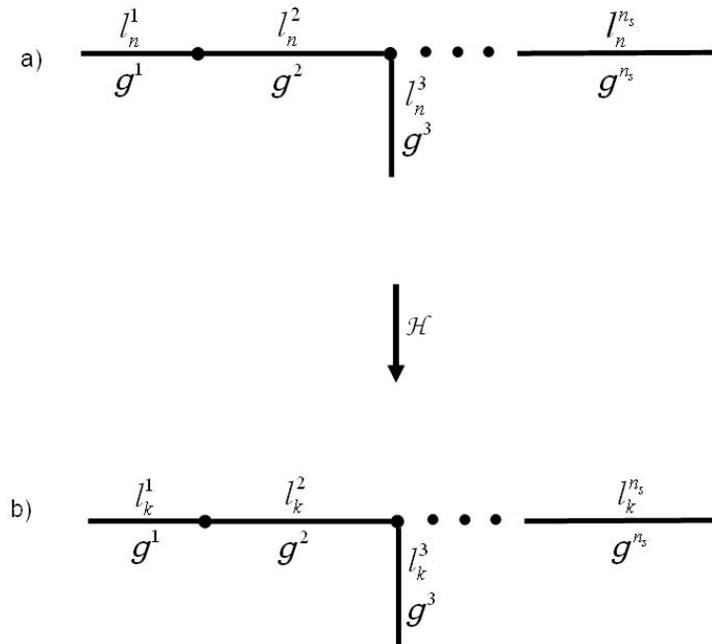


Figura 4.6: a) Enlace obtido pelo AG na n -ésima geração e que serve de estimativa inicial a técnica local de otimização b) Enlace após a k -ésima execução do bloco \mathcal{H} , onde $k = n/N_{\mathcal{H}}$. Apenas o comprimento de cada seção é otimizado pela técnica de otimização local.

que passa a ser usado como informação *a priori* na próxima execução do AG. O valor s é incrementado para que a próxima seção possa ser testada pelo processo de realimentação.

As seções armazenadas em Γ têm suas estruturas topológicas, exceto as bitolas, fixas até o final do processo de otimização; os comprimentos e bitolas continuam sendo otimizados. Entretanto, a otimização é restrita a uma pequena faixa em torno do valor do comprimento armazenado em Γ ou a bitolas que estão contíguas à bitola armazenada em Γ , atuando dessa forma como um processo de ajuste fino desses parâmetros. No caso da bitola ser, por exemplo, 0,5mm, no processo de decodificação apenas as bitolas [0,4 0,5 0,63]mm podem ser escolhidas para essa seção em Γ . Já para um comprimento l , a otimização seria realizada na faixa de $l - l\Delta_L$ a $l + l\Delta_L$. O valor de Δ_L pode variar de 0,05 a 0,20, sendo definido pelo usuário antes do processo de otimização.

Se o comando para inicializar o AG não ocorrer, a condições de parada são verificadas. No entanto, nesse caso, são testadas as melhores soluções atuais obtidas pela técnica local, isto é, se $V_H(\tilde{\Theta}_{S_{11}}^k) + V_{S_{11}}(\tilde{\Theta}_{S_{11}}^k) < \delta_s$, o processo de identificação é finalizado.

4.2.3 Processo de Codificação e Decodificação

O processo de codificação/decodificação é um passo importante no processo de otimização por AG. Basicamente, esse processo conecta o AG ao problema físico sob otimização e pode determinar a eficiência da pesquisa no espaço de soluções, reduzindo a área de pesquisa por ignorar soluções não factíveis ou aquelas que não aplicáveis ao problema em otimização. As estratégias de codificação/decodificação propostas são descritas abaixo.

Neste trabalho, todos os genes são codificados como números reais que variam de 0 a 1. Cada cromossomo tem genes suficientes para armazenar informação sobre um enlace com o número máximo de seções N_s , ainda que o número real de seções n_s seja menor. Mais especificamente, cada cromossomo é composto por $3N_s + 1$ genes, sendo: $2N_s$ genes utilizados para determinar o comprimento e a bitola de cada seção; $N_s - 1$ para determinar os tipos (serial ou *bridged-tap*) das seções do enlace, com exceção da primeira seção que é sempre serial; e dois genes associados ao número de seções n_s e ao número de *bridged-taps* n_b . Como $n_s \leq N_s$, o número de genes não é maior que o comprimento do vetor de parâmetros Θ , $|\Theta|$, e há uma correspondência de cada elemento de Θ a um gene (desconsiderando o gene que seria necessário para codificar o tipo da primeira seção). Na Figura 4.7 é mostrado um exemplo de codificação de um enlace DSL mostrado em (a) em seu cromossomo correspondente mostrado em (b). Nesse caso, o número máximo de seções N_s é igual a 4 e assim o número de genes é igual a 13. A função de cada gene mostrado será descrita nas próximas seções. O processo de codificação \mathcal{G} mapeia os parâmetros do enlace no seu cromossomo codificado. Por outro lado, o processo de decodificação \mathcal{G}^{-1} transforma o conjunto de número reais do cromossomo em um enlace correspondente. O processo de codificação \mathcal{G} (ou decodificação \mathcal{G}^{-1}) pode ser aplicado apenas a um determinado parâmetro; por exemplo, o valor codificado do comprimento da segunda seção l^2 é representado por $\mathcal{G}(l^2)$.

4.2.3.1 Codificação/decodificação do Comprimento e Bitola

A representação do comprimento da s -ésima seção é dada por:

$$\mathcal{G}(l^s) = (l^s - L_{\min}) / (L_{\max} - L_{\min}), \quad (4.5)$$

onde L_{\min} e L_{\max} representa o limite mínimo e máximo do comprimento definido pelo usuário, respectivamente. Essa expressão assegura que o comprimento de uma seção sempre estará dentro de uma faixa factível. Os valores de L_{\min} e L_{\max} são diferentes dependendo se a seção é serial ou *bridged-tap*.

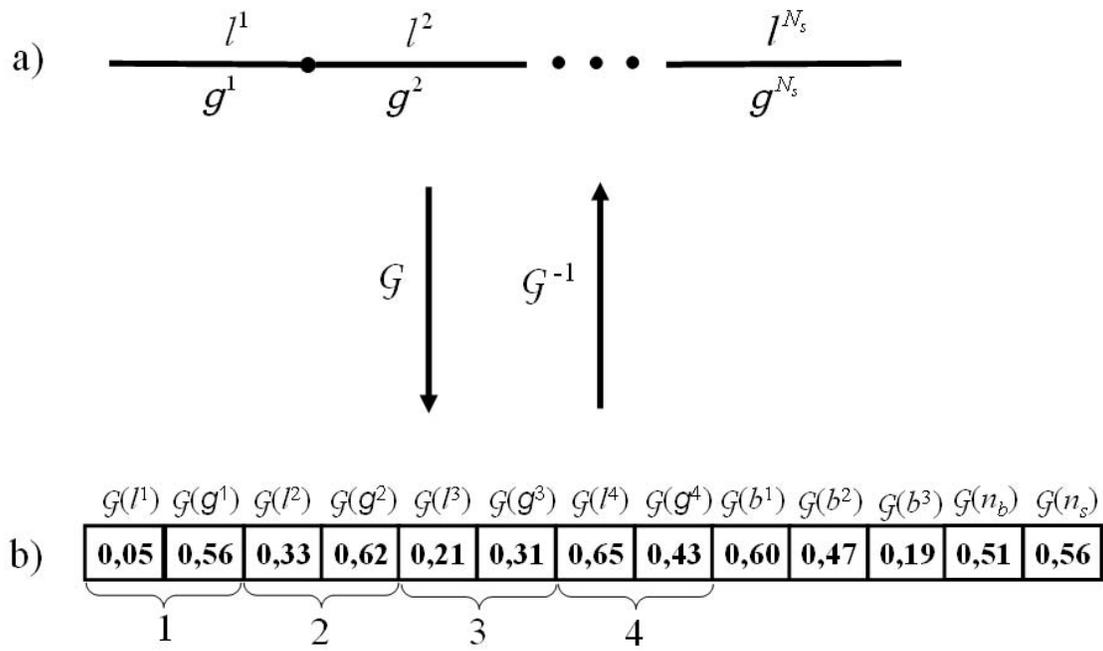


Figura 4.7: Processo de codificação (\mathcal{G}) e decodificação (\mathcal{G}^{-1}) para um enlace com o número máximo de seções N_s igual a 4, cada uma com comprimento l^i e bitola g^i .

De forma similar, a representação da bitola da s -ésima seção é dada por:

$$\mathcal{G}(g^s) = (g^s - G_{\min}) / (G_{\max} - G_{\min}). \quad (4.6)$$

Como na prática as bitolas assumem valores discretos, sendo definidas por um vetor $[G_1, \dots, G_M]$ com M elementos, uma representação mais adequada seria levar em consideração essa restrição. Por exemplo, as possíveis bitolas de enlaces ETSI são $[0,32 \ 0,40 \ 0,50 \ 0,63 \ 0,90]$ mm. Assim, três representações alternativas foram consideradas.

A primeira, denominada *freegauge*, é a mais simples e consiste em mapear a faixa de $[0,1]$ em M intervalos, com cada intervalo correspondendo a uma possível bitola (por exemplo, $M = 5$ para os enlaces ETSI). Dado um valor de gene $\mathcal{G}(g^s)$ e assumindo-se que as bitolas disponíveis estão organizadas em um vetor, a posição da bitola correspondendo ao gene g^s no vetor, é dada por:

$$j = \lceil M\mathcal{G}(g^s) \rceil; \quad (4.7)$$

isto é, a bitola referente ao gene g^s é o j -ésimo elemento do vetor; onde $\lceil \cdot \rceil$ é a função *ceil*. Por conveniência, esta operação será denotada pela expressão:

$$[y, j] = \mathcal{U}(z, [G_1, \dots, G_M]) \quad (4.8)$$

onde \mathcal{U} é uma função que retorna dois valores, y e j . O valor y é um elemento do vetor $[G_1, \dots, G_M]$ e j é o seu índice, como na expressão 4.7. O valor de entrada $z \in [0, 1]$ seleciona

o j -ésimo elemento do vetor $[G_1, \dots, G_M]$. Em outras palavras, a estratégia *freegauge* pode ser representada por:

$$[g^s, j] = \mathcal{U}(\mathcal{G}(g^s), [G_1, \dots, G_M]). \quad (4.9)$$

Uma estratégia mais elaborada pode ser usada considerando-se que os enlaces não apresentam repetição de bitolas e que as bitolas estão em ordem crescente. Esse é um caso típico em enlaces com apenas seções seriais descritos nas recomendações ITU-T [58]. Nessa estratégia, chamada *gaugesort*, o número de possíveis sub-intervalos associados a cada seção é dependente não somente do número de elementos M do vetor $[G_1, \dots, G_M]$, mas também do número de seções do enlace e das bitolas das seções anteriores. Na estratégia *gaugesort*, todos os genes $\{\mathcal{G}(g^1), \dots, \mathcal{G}(g^{n_s})\}$ correspondendo as bitolas $\{g^1, \dots, g^{n_s}\} \in \Theta$ são decodificados juntos. O código para esta técnica é mostrado na Tabela 4.2, onde é assumido que os elementos de $[G_1, \dots, G_M]$ estão em ordem crescente.

Tabela 4.2: A técnica *gaugesort* para a decodificação dos valores de bitolas.

Inputs: $n_s, \{\mathcal{G}(g^1), \dots, \mathcal{G}(g^{n_s})\}, [G_1, \dots, G_M]$

Output: $\{g^1, \dots, g^{n_s}\}$

1. $j = 0$;
2. **For** $s = 1$ to n_s ,
3. $S = M - j - n_s + s$;
4. $[g^s, j'] = \mathcal{U}(\mathcal{G}(g^s), [G_{j+1}, G_{j+2}, \dots, G_{j+S}]);$
5. $j = j + j'$;
6. **End**

O valor de j da Tabela 4.2 determina o índice da última bitola usada do vetor $[G_1, \dots, G_M]$. Basicamente, o número de bitolas disponíveis utilizadas pela função \mathcal{U} para definir a bitola de uma dada seção é igual ao número de bitolas ainda não utilizadas ($M - j$) menos o número de seções ainda não decodificadas ($n_s - s$). Deste modo, a técnica evita a repetição de bitolas e assegura que sempre estarão em ordem crescente a partir do CO. Para aplicação desse método, deve-se assegurar que o número de bitolas seja pelo menos igual ao número de seções.

O terceiro método é mais geral que o método *gaugesort*, porém mais restritivo que o método *freegauge*. Esse método foi denominado de *typegauge*. No método *typegauge* duas seções podem apresentar o mesmo valor de bitola contanto que sejam de tipos diferentes, *bridged-tap* - serial e serial - *bridged-tap*; essa é uma situação comum nos enlaces descritos nas recomendações [58]. Basicamente se a seção for do tipo *bridged-tap*, qualquer bitola pode ser escolhida do vetor $[G_1, G_2, \dots, G_M]$, e assim a bitola é determinada pela expressão

$$[g^s, \text{ndx}] = \mathcal{U}(\mathcal{G}(g^s), [G_1, G_2, \dots, G_M]). \quad (4.10)$$

O mesmo ocorre se o tipo da seção anterior à seção que está sendo decodificada também for *bridged-tap*. No entanto no caso de duas seções seriais consecutivas, a bitola da seção anterior $g^{s-1} = G_m$ é retirada do vetor de bitolas $[G_1, G_2, \dots, G_{m-1}, G_m, G_{m+1}, \dots, G_M]$ e após isso é aplicada a expressão

$$[g^s, \text{ndx}] = \mathcal{U}(\mathcal{G}(g^s), [G_1, G_2, \dots, G_{m-1}, G_{m+1}, \dots, G_M]), \quad (4.11)$$

para determinar a bitola da seção atual g^s . Assim duas seções seriais nunca apresentarão a mesma bitola como pode acontecer no método *freegauge*. Além disso, pelas próprias características do método *typegauge*, é possível evitar que algumas bitolas ocorram em seções nas quais elas não são encontradas na prática. Um exemplo é a bitola 0,32 mm, que apresenta alta atenuação, e por isso é instalada apenas próximo ao CO e com comprimento de poucos metros. Usando o método *typegauge* é possível facilmente forçar que essa bitola ocorra apenas na primeira seção do enlace, como ocorre na prática.

4.2.3.2 Decodificação do Número de Seções

O valor de n_s está relacionado ao seu gene pela expressão a seguir:

$$n_s = \text{round} [N_s^{\min} - 0,5 + (N_s - N_s^{\min} + 1)\mathcal{G}(n_s)], \quad (4.12)$$

onde N_s^{\min} é o número mínimo de seções e a função **round** arredonda um número para um valor inteiro mais próximo. As constantes 0,5 e 1 foram inseridas na Equação 4.12 para garantir que os valores extremos N_s^{\min} e N_s tivessem a mesma probabilidade de ocorrência que os valores intermediários entre eles. Tanto o número mínimo de seções N_s^{\min} quanto o número máximo de seções N_s são configurados pelo usuário. Se houver informação *a priori* sobre o número de seções do enlace, então o usuário deve configurar $N_s^{\min} = N_s$, e assim o número de seções n_s será sempre igual a N_s^{\min} .

4.2.3.3 Decodificação do Tipo das Seções

Decodificar o tipo das seções é equivalente a identificar quais seções são *bridged-taps*. O procedimento começa pela determinação do número n_b de *bridged-taps* e é semelhante ao processo de decodificação do número de seções. No entanto, em vez de ser configurado pelo usuário, o número de máximo de *bridged-taps* N_b é determinado pelo número de seções n_s por:

$$N_b = \text{round} [(n_s - 1) / 2]. \quad (4.13)$$

Essa expressão advém do fato que as *bridged-taps* devem ser alternadas com as seções seriais ao longo do enlace e a primeira seção ser sempre serial.

A partir do número máximo de *bridged-taps* N_b , as relações entre o número de *bridged-taps* n_b e seu gene codificado é dada por:

$$n_b = \text{round} [N_b^{\min} - 0,5 + (N_b - N_b^{\min} + 1)\mathcal{G}(n_b)], \quad (4.14)$$

onde N_b^{\min} é o número mínimo de *bridged-taps* configurado pelo usuário.

A posição das *bridged-taps* é determinada por um vetor de genes $[\mathcal{G}(b^1), \dots, \mathcal{G}(b^{N_s-1})]$ (mostrado na Figura 4.7) com $N_s - 1$ elementos no qual apenas os primeiros $n_s - 1$ são relevantes. O gene $\mathcal{G}(b^s)$ pertencente ao cromossomo $\mathcal{G}(\Theta)$, onde s varia de 1 a $n_s - 1$, pode ser visto como um valor booleano que indica se a $(s + 1)$ -ésima seção é uma *bridged-tap*. No entanto, devido a alternância entre seções seriais e *bridged-taps*, não é possível determinar a posição de uma *bridged-tap* considerando apenas o valor de um gene, sendo o tipo de uma seção dependente do tipo das outras seções; com exceção de enlaces que apresentem apenas uma *bridged-tap*, pois nesse caso não existe a possibilidade de ocorrência de duas *bridged-taps* consecutivas. Assim, a posição da *bridged-tap* é escolhida independentemente do tipo das outras seções. Semelhante ao processo *gaugesort*, a operação de decodificação \mathcal{G}^{-1} da posição das *bridged-taps* é executada considerando o conjunto de genes associados.

A decodificação é feita iterativamente verificando se o valor do s -ésimo gene é o maior entre os genes do vetor $[\mathcal{G}(b^1), \mathcal{G}(b^2) \dots, \mathcal{G}(b^{n_s-1})]$, conforme mostrado na Tabela 4.3. A função *indexofmax* retorna o índice do valor máximo do vetor de entrada. Assume-se que a variável h_n armazena a posição da n -ésima *bridged-tap* previamente decodificada. Por exemplo, $n_s = 6$ e $h_2 = 4$ significa que para um enlace de 6 seções, já existem pelo menos duas *bridged-taps* e a segunda *bridged-tap* está localizada na quarta seção do enlace. Nesse caso, a quinta seção é obrigatoriamente serial e deve-se verificar apenas se a sexta seção é uma *bridged-tap* ou uma seção serial. Após armazenada a posição da n -ésima *bridged-tap*, seu gene correspondente é inicializado com zero ($\mathcal{G}(b^{(h_n-1)}) = 0$), assim como os genes adjacentes a ele ($\mathcal{G}(b^{(h_n-2)}) = 0$ e $\mathcal{G}(b^{(h_n)}) = 0$). Dessa forma, garante-se que não haverá *bridged-taps* consecutivas. O processo continua até que todas as *bridged-taps* tenham suas posições determinadas.

4.2.3.4 Um Exemplo de Decodificação

Esse exemplo mostra como obter Θ para o cromossomo mostrado na Figura 4.8. Nesse caso, é assumido número mínimo de seções $N_s^{\min} = 1$, número máximo de seções $N_s = 4$ e número mínimo de *bridged-taps* $N_b^{\min} = 0$. Primeiro, o número de seções é decodificado usando o último gene $\mathcal{G}(n_s) = 0,56$, obtendo-se $n_s = \text{round} [4 - 0,5 + (4 - 1 + 1)0,56] = 3$.

Tabela 4.3: Decodificação da posição das *bridged-taps*.**Inputs:** $n_s, n_b, [\mathcal{G}(b^1), \dots, \mathcal{G}(b^{N_s-1})]$ **Output:** $\{b^1, \dots, b^{n_s}\}$

1. For $n = 1$ to n_b ,
2. $h_n = 1 + \text{indexofmax}[\mathcal{G}(b^1), \mathcal{G}(b^2), \dots, \mathcal{G}(b^{(n_s-1)})];$
3. $\mathcal{G}(b^{(h_n-1)}) = 0, \mathcal{G}(b^{(h_n-2)}) = 0$ e $\mathcal{G}(b^{(h_n)}) = 0$
4. End

$\mathcal{G}(l^1)$	$\mathcal{G}(g^1)$	$\mathcal{G}(l^2)$	$\mathcal{G}(g^2)$	$\mathcal{G}(l^3)$	$\mathcal{G}(g^3)$	$\mathcal{G}(l^4)$	$\mathcal{G}(g^4)$	$\mathcal{G}(b^1)$	$\mathcal{G}(b^2)$	$\mathcal{G}(b^3)$	$\mathcal{G}(n_b)$	$\mathcal{G}(n_s)$
0,05	0,56	0,33	0,62	0,21	0,31	0,65	0,43	0,60	0,47	0,19	0,51	0,56
1		2		3		4						

Figura 4.8: Exemplo de um cromossomo para ilustrar o processo de decodificação. Nesse caso o número máximo de seções N_s é igual a 4.

Baseado no valor de n_s , do gene $\mathcal{G}(n_b)$ e usando as equações (4.13) e (4.14), pode-se obter o valor do número de *bridged-taps* n_b , que nesse caso é igual 1. A posição da *bridged-tap* é determinada por: $h_1 = 1 + \text{indexofmax}[\mathcal{G}(b^1), \mathcal{G}(b^2)] = 1 + \text{indexofmax}[0,60, 0,47]$. Nesse caso, como a função *indexofmax* retorna 1, $h_1 = 2$ e assim a *bridged-tap* está localizada na segunda seção. Para a decodificação dos comprimentos e das bitolas das seções são usados apenas os genes $\{\mathcal{G}(l^1), \dots, \mathcal{G}(l^3)\}$ e $\{\mathcal{G}(g^1), \dots, \mathcal{G}(g^3)\}$ correspondendo aos genes das três primeiras seções, respectivamente. Considerando $L_{\min} = 100$ m e $L_{\max} = 4000$ m, o comprimento da primeira seção, calculada pela Equação 4.5, é $l^1 = 300$ m já que $\mathcal{G}(l^1) = 0,05$. Usando o método *freegauge*, dado pela Equação 4.9, e assumindo que as possíveis bitolas são $[0,32, 0,40, 0,50, 0,63, 0,9]$ mm, a bitola da primeira seção é $g^1 = 0,5$ mm dado que $\mathcal{G}(g^1) = 0,56$. Raciocínio similar aplica-se às outras seções.

4.2.4 Controle Auto-adaptativo dos Parâmetros do AG

Essencialmente, existem dois parâmetros de controle do processo do AG: probabilidade de cruzamento (p_c) e a probabilidade de mutação (p_m). A probabilidade p_c controla o processo de cruzamento e determina se dois indivíduos selecionados trocarão suas informações genéticas entre si para gerar dois novos indivíduos. A probabilidade p_m controla o processo de mutação que é aplicado a cada gene de um cromossomo, determinando se o gene sofrerá mutação. Quando um gene $\mathcal{G}(x)$ é selecionado para sofrer mutação, seu novo valor $\mathcal{G}(x')$ é determinado

por:

$$\mathcal{G}(x') = \mathcal{G}(x) + \mathcal{N}(0, \sigma_m),$$

onde $\mathcal{N}(0, \sigma_m)$ denota uma distribuição normal com média μ igual a zero e desvio padrão σ_m .

Uma regra geral, difundida na literatura, é usar valores fixos para todos os parâmetros do AG, descrito pelo vetor $\Omega = [p_c, p_m, \sigma_m]$, que também inclui o parâmetro σ_m , por ter também bastante impacto na diversidade da população. Além disso, o número total de gerações N do AG é configurado antes de sua execução e mantém o mesmo valor ao longo de todo o processo de otimização. Mas, em muitas aplicações, este procedimento pode ser ineficiente e o algoritmo pode ficar preso à mínimos locais. Por esse motivo, uma heurística de controle auto-adaptativa foi desenvolvida para determinar os valores ótimos para os parâmetros do AG em Ω e também o valor de N ao longo do processo de otimização.

A estratégia adotada basicamente consiste em variar os valores associados a Ω a partir de valores altos (em torno de 0,90) até valores bem pequenos (em torno de 0,1). Essa variação é condicionada ao aumento do valor de aptidão do melhor indivíduo do AG, ou em outras palavras, a evolução do melhor indivíduo do AG. Um conjunto de valores dos parâmetros do AG que faz o melhor indivíduo evoluir é mantido constante até que não haja mais evolução desse indivíduo. Quando essa condição é detectada, os parâmetros do AG têm seus valores reduzidos. Adicionalmente, sempre que qualquer evolução do melhor indivíduo aconteça, o número total de gerações N é aumentado, dando mais tempo ao algoritmo para continuar a sua pesquisa. A idéia por trás dessa estratégia, é que no princípio do processo de otimização, o AG tem pouca informação sobre o espaço de solução e assim deve ser mantida o máximo possível à diversidade da população. Quanto maior o valor dos parâmetros do AG maior será a diversidade de sua população. Por outro lado, ao final do processo de otimização, o AG tem bastante informação sobre o espaço de solução e as melhores soluções encontradas necessitam apenas serem refinadas. Portanto, nesse caso, é desejável que os parâmetros do AG tenham valores baixos para manter a sua pesquisa restrita em torno dessas possíveis melhores soluções. Como a variação nos valores dos parâmetros do AG está relacionada à evolução do melhor indivíduo, os valores finais desses parâmetros podem variar de uma simulação para outra.

O código que implementa esse controle de parâmetros auto-adaptativo é mostrado em mais detalhes na Tabela 4.4, onde p_c^n , p_m^n e σ_m^n são os valores da probabilidade de cruzamento, probabilidade de mutação e do desvio padrão na geração n , respectivamente; P_c , P_m e Σ_m são respectivamente os valores máximos da probabilidade de cruzamento, probabilidade de mutação e do desvio padrão, e são configurados pelo usuário antes do processo de otimização. c é um contador que armazena o número de gerações sucessivas sem evolução e os parâmetros δ_{p_c} , δ_{p_m} e δ_{σ_m} indicam a taxa de decaimento de p_c , p_m e σ_m , respectivamente, e também são configurados pelo usuário.

A cada geração n , depois de calcular a matriz de objetivos \mathbf{Q}^n da população Φ^n , os melhores indivíduos $\tilde{\Theta}_H^n$ e $\tilde{\Theta}_{S_{11}}^n$ dessa população tem os seus objetivos, $V_H(\tilde{\Theta}_H^n)$ e $V_{S_{11}}(\tilde{\Theta}_{S_{11}}^n)$, comparados com objetivos dos melhores indivíduos em todas as gerações $V_H(\tilde{\Theta}_H)$ e $V_{S_{11}}(\tilde{\Theta}_{S_{11}})$, respectivamente. Se qualquer evolução em pelo menos uma dessas melhores soluções acontece, isto é, se $V_H(\tilde{\Theta}_H^n) < V_H(\tilde{\Theta}_H)$ ou $V_{S_{11}}(\tilde{\Theta}_{S_{11}}^n) < V_{S_{11}}(\tilde{\Theta}_{S_{11}})$, c é levado a 0 e N é aumentado por ΔN . Caso contrário, c é incrementado de 1. Adicionalmente, c é avaliado pela expressão $w(N-n)$, onde w (em torno de 0.05-0.1) é configurado pelo usuário. Se c é maior que $w(N-n)$, os valores atuais dos parâmetros do AG são reduzidos de acordo com a Tabela 4.4, e c é levado a 0.

A expressão $w(N-n)$ também controla a taxa de variação dos parâmetros do AG. No princípio do processo de otimização, os parâmetros do AG variam lentamente devido a N ser muito maior que n . Por outro lado, ao término do processo de otimização, os parâmetros do AG são reduzidos mais rapidamente devido a n estar mais próximo a N .

Tabela 4.4: Controle auto-adaptativo dos parâmetros do AG.

Inputs: $p_c^{n-1}, p_m^{n-1}, \sigma_m^{n-1}, \delta_{p_c}, \delta_{p_m}, \delta_{\sigma_m}$

Outputs: p_c^n, p_m^n, σ_m^n

1.	For $n = 1, \dots, N$
2.	$\mathbf{Q}^n = \text{costfunction}(\Phi^n);$
3.	$V_H(\tilde{\Theta}_H^n) = \min_{r=1, \dots, R} \tilde{\Theta}_H^n;$
4.	$V_{S_{11}}(\tilde{\Theta}_{S_{11}}^n) = \min_{r=1, \dots, R} \tilde{\Theta}_{S_{11}}^n;$
5.	If ($V_H(\tilde{\Theta}_H^n) < V_H(\tilde{\Theta}_H)$ or $V_{S_{11}}(\tilde{\Theta}_{S_{11}}^n) < V_{S_{11}}(\tilde{\Theta}_{S_{11}})$),
6.	$N = N + \Delta N;$
7.	$c = 0;$
8.	Else
9.	$c = c + 1;$
10.	End
11.	If $c > w(N-n),$
12.	$p_m^n = p_m^{n-1} - p_m^{n-1} \delta_{p_m};$
13.	$p_c^n = p_c^{n-1} - p_c^{n-1} \delta_{p_c};$
14.	$\sigma_m^n = \sigma_m^{n-1} - \sigma_m^{n-1} \delta_{\sigma_m};$
15.	$c = 0;$
16.	End
17.	End

4.3 Ferramenta de Medição Automática

Pode-se dizer que o processo de identificação do enlace começa bem antes da aplicação da metodologia DELTI. É iniciada durante o processo de medição das quantidades de interesse ou curvas alvo, função de transferência H e do parâmetro de espalhamento S_{11} , usando os setups descritos no Apêndice A. A obtenção de medidas confiáveis e precisas é de grande importância para a precisão dos resultados obtidos por DELTI. Geralmente no processo de medição, é realizada a medida de uma mesma curva alvo diversas vezes e em seguida realizado o seu tratamento estatístico. Esse procedimento pode ser complexo dependendo do número de repetições. A redução da intervenção humana durante esse processo pode minimizar a ocorrência de erros que poderiam comprometer a precisão dos dados medidos. Por esse motivo, foi desenvolvida uma rotina de medição, chamada de AMT (*Automatic Measurement Tool*), que além de controlar automaticamente o processo de medição, faz a análise estatística dos dados medidos, obtendo o seu intervalo de confiança e o seu nível de discrepâncias, denominados de *outliers*, que são eliminados durante a aplicação desse processo. A AMT ainda necessita ser estendida para tratar dados medidos provenientes de modems e do DSLAM, que são menos precisos que os dos equipamentos dedicados de medição e apresentam certas particularidades, como por exemplo, a ação de filtros que eliminam partes da banda de frequência ou problemas devido a calibração não-uniforme ao longo da banda de medição. Na Figura 4.9 é mostrado o diagrama de bloco da AMT.

Inicialmente, são configurados pelo usuário os parâmetros relacionados à medição como o número de repetições, a faixa de frequência, a topologia do enlace, aspectos relacionados à calibração e o parâmetro a ser medido (função de transferência H ou parâmetro de espalhamento S_{11}). Iniciado o processo de medição, a obtenção dos dados medidos e seu armazenamento é realizado de forma automática. À seguir, é feita a análise estatística dos dados obtidos, sendo avaliada a precisão do processo de medição, verificando-se o número de *outliers* encontrados, o qual se ultrapassa um patamar pré-estabelecido, determina a reinicialização do processo de medição. Caso contrário, os tons de frequências em que se encontram os *outliers* são excluídos dos dados medidos. Por fim, as curvas medidas e suas respectivas estatísticas são armazenadas em um meio permanente de armazenamento (disco rígido, por exemplo). A ferramenta de medição permite também a comparação dos dados medidos a modelos de linha na forma de tabelas, arquivos de texto e gráficos.

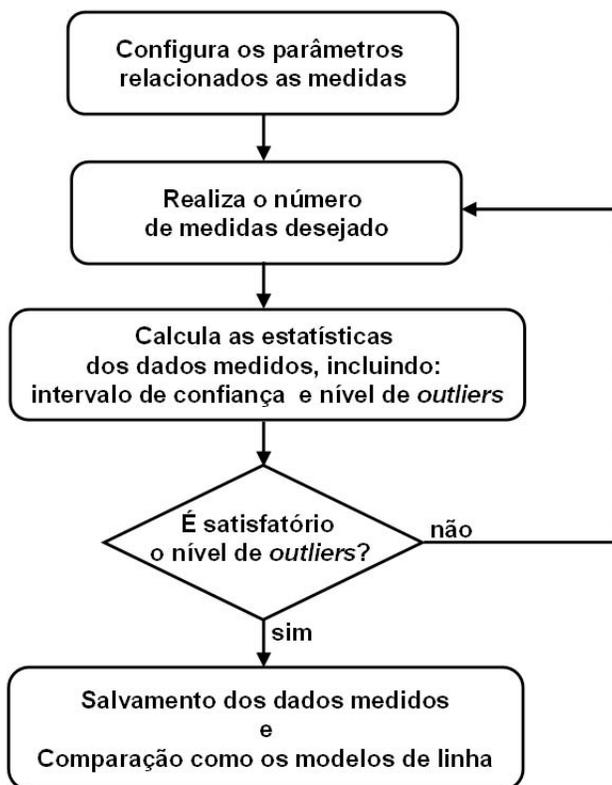


Figura 4.9: Diagrama de bloco da AMT.

4.4 Conclusões do Capítulo

Neste capítulo foi discutido o método TIMEC que usa algoritmos genéticos multiobjetivos e técnicas locais de otimização. O TIMEC é a principal parte da metodologia de identificação de enlaces LQ-SUITE. Como se pôde perceber o método foi especializado considerando as características do problema de identificação de enlaces DSL. Dessa forma, sob alguns aspectos, o método perdeu sua generalidade em prol de maior desempenho e redução do custo computacional. A interação das técnicas de otimização global, baseada em AG, e a técnica local, cada uma trabalhando em nichos diferentes, aumentou a precisão do método e reduziu seu tempo de processamento. Além disso, as técnicas de codificação\decodificação mostraram que TIMEC é flexível, aceitando diversos métodos baseados na quantidade e qualidade de informações pré-existentes sobre o enlace. Dessa forma, quanto mais conhecimento se tiver da rede telefônica, maior a quantidade de inteligência que será possível de aplicar ao TIMEC, de forma rápida e simples. Um dos aspectos que podem prejudicar a utilização de algoritmos genéticos a problemas práticos, é o ajuste de seus parâmetros internos, como a probabilidade de cruzamento e mutação. Por isso, nessa tese foi desenvolvido um método que ajusta esses parâmetros automaticamente dependendo das características de evolução do al-

goritmo genético. Assim, na prática o ajuste desses parâmetros será transparente aos técnicos responsáveis de operar a metodologia de identificação. Outra automatização aplicada foi o desenvolvimento de uma ferramenta de medição que diminuí a interferência humana, realizando controle dos equipamentos de medidas, a coleta dos dados, seu tratamento e gravação dos dados obtidos em disco.

No próximo capítulo serão apresentadas as técnicas de pré-processamento que são utilizadas para extrair informações sobre o enlace, que são utilizadas pelo TIMEC. Essas técnicas se baseiam na análise no tempo de medidas de uma porta e de duas portas.

Capítulo 5

Técnicas Analíticas de Pré-processamento

Neste capítulo são descritos dois módulos de pré-processamento que são utilizados para obter informação *a priori* sobre um enlace DSL. O primeiro deles, chamado IRA, é baseado na análise da resposta impulsiva, no qual são extraídas informações como o comprimento total e o número de *bridged-taps* existentes no enlace. O segundo, chamado de FSF, é baseado na análise da resposta TDR do enlace, podendo obter informação do comprimento total e da topologia da primeira seção do enlace. Finalmente, na última seção são apresentadas as conclusões.

5.1 Técnicas de Pré-processamento para Obtenção de Informação *a priori*

Para melhorar o desempenho e a confiabilidade da metodologia TIMEC, pode-se utilizar informações *a priori* sobre o enlace, que podem incluir desde o seu comprimento total, o número de *bridged-taps*, a topologia de uma seção em particular, entre outras. Nesse trabalho, duas técnicas foram desenvolvidas para obtenção de informação sobre o enlace. A primeira trata da análise da resposta impulsiva do enlace. Com o advento dos modems ADSL2 e ADSL2+, é possível medir diretamente a função de transferência do enlace (medida de DELT), e da transformada de Fourier inversa dessa medida obter a sua resposta impulsiva, utilizada pela técnica de análise desenvolvida. A essa técnica foi dado o nome de IRA (em inglês, *Impulse Response Analysis*), e é capaz de obter o comprimento total do enlace, o número de *bridged-*

taps e os comprimentos dessas *bridged-taps*. A outra técnica que tem o nome de FSF (em inglês *First Section Finder*) baseia-se na análise do reflectograma TDR, que é obtido da impedância de entrada do enlace. Essa técnica utiliza uma pequena parte do método descrito em [11], referente a estimação das bitolas das seções. A maior parte do algoritmo se baseia na interpretação das informações obtidas por um extrator de característica chamado WFE (*Wavelet-Based Feature Extraction*) [59]. Com o FSF é possível obter a estimação da primeira seção do enlace e para enlaces com *bridged-taps*, o número de *bridged-taps* e informação sobre a posição dessas *bridged-taps*. Cada uma dessas técnicas será descrita a seguir. Mas antes, será realizada uma breve descrição do WFE, que é uma ferramenta que auxilia ambas as técnicas nos seus processos de estimação.

5.1.1 Extrator de Características baseado em Wavelet - WFE

Uma resposta TDR e outros sinais no tempo de um enlace são compostos por um conjunto de ecos (ou picos) que podem ser classificados como ecos reais e espúrios. Um eco real é a primeira reflexão de uma descontinuidade do enlace e, teoricamente, contém toda a informação necessária para detectar, localizar e identificar a descontinuidade que a gerou. Enquanto ecos espúrios são definidos como todos os ecos causados por sucessivas reflexões e não indicam a presença de descontinuidade. Cada eco, assim como o sinal de entrada, tem uma borda de subida e descida cujo intervalo de tempo definido entre elas é aproximadamente o mesmo que a largura do pulso de entrada. Assim, o intervalo do eco pode ser definido como o intervalo entre o tempo de chegada de um dado eco na porta de entrada do enlace (borda de subida) e o tempo que o eco começa a dissipar (borda de descida).

WFE (*Wavelet-Based Feature Extraction*) é um método geral para extração de características que pode ser aplicado a medidas de SELT e DELT usualmente disponíveis em aplicações DSL [59]. O método baseia-se na análise das singularidades do sinal, como as bordas de subida e descida, detectadas através da transformada wavelet, por um conjunto de regras definidas para cada tipo de sinal. Especificamente, o WFE é capaz de extrair as características do sinais no domínio do tempo como a resposta impulsiva, a resposta TDR, e as versões no domínio do tempo do parâmetro de espalhamento e da impedância de entrada.

O método WFE consiste de duas etapas: uma de pré-processamento e outra de identificação das características do sinal. A primeira etapa segue os seguintes passos [59]:

1. Cálculo da Transformada Wavelet Contínua (TWC)
2. Obtenção dos extremos locais da TWC considerando a seguinte equação de limiar

$$T_i = k_s \left[\frac{\sum_{j=1}^n |Wf(u_j, s_i)|}{n} + \sigma_{u, s_i}^2 \right], \quad (5.1)$$

onde $Wf(u, s)$ é o TWC do sinal f ; s_i é i -ésima escala; n é o número dos elementos na i -ésima escala, que é igual ao número de amostras do sinal f , σ_{u, s_i}^2 é a variância dos coeficientes da escala s_i . A constante de sensibilidade k_s é usada para permitir ao usuário a opção de aumentar ou diminuir a sensibilidade do algoritmo na pesquisa por extremos diminuindo e aumentando o valor de k_s , respectivamente. Configurando $k_s = 0.5$ significa duplicar a sensibilidade do algoritmo e $k_s = 2$ significa diminuir pela metade a sensibilidade do algoritmo.

3. A partir dos extremos locais da CWT, o algoritmo encadeia os extremos em linhas. A partir da convergência dessas linhas até a escala mais "fina", o algoritmo encontra as singularidades presentes no sinal. O segundo bloco do algoritmo toma essas singularidades e as submete a um conjunto de regras de acordo com o tipo de sinal e obtém as características desejadas, como a localização de ecos no reflectograma TDR, por exemplo.

Como sinais usados em sistemas DSL podem ter características bem diferentes, considerando, por exemplo, o número de amostras do sinal, a constante de sensibilidade k_s tem de ser diferente para cada sinal. Especificamente, $h(t)$ e $s_{11}(t)$ tem frequência de amostragem e assim número de pontos de amostras maior que as curvas TDR, além das singularidades ocorrerem mais próximas umas das outras e terem mais intensidade para esses sinais [60], sendo nesse caso necessário aumentar a sensibilidade da detecção de máximos na escala wavelet.

Ainda para um mesmo sinal, a escolha apropriada de k_s pode não ser trivial, variando devido ao comprimento total e as características da linha. Assim foi desenvolvido um método que determina o valor de k_s para uma resposta TDR em particular. Como mostrado na Figura 5.1, o número de pontos e quais pontos são determinados depende do valor de k_s . Dessa forma, diferentes informações podem ser obtidas, com acréscimo de informações incorretas devido a ecos espúrios detectados como ecos verdadeiros ou informações perdidas devido a ecos verdadeiros não serem detectados. O critério adotado leva em conta que a parte do sinal mais característica é o pico de final de enlace. Assim, o valor de k_s é variado de um valor alto (baixa sensibilidade) a um valor baixo (alta sensibilidade), até que ambas as bordas de subida e descida do pico de final de enlace sejam encontradas, quando o valor de k_s pode ser considerado determinado. Assim, pelo menos a detecção desse pico, definido pela sua borda de subida e descida, é garantida. Além disso, evita-se utilizar valores muito elevados de k_s , que eliminariam todos os espúrios mas com a eventual perda de informações.

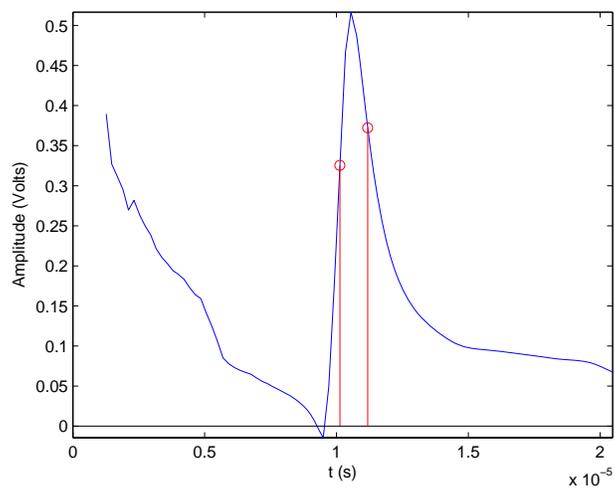
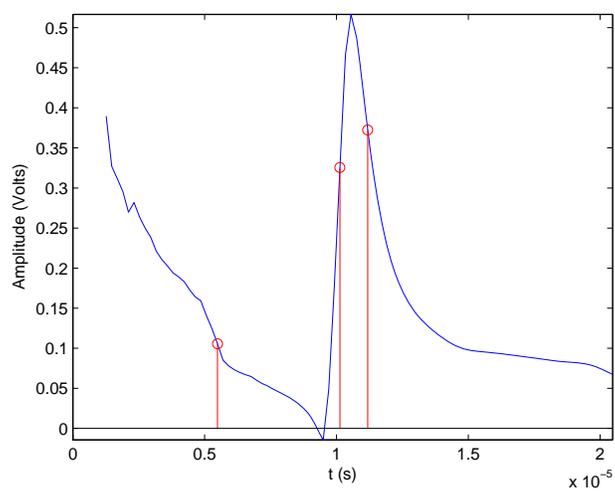
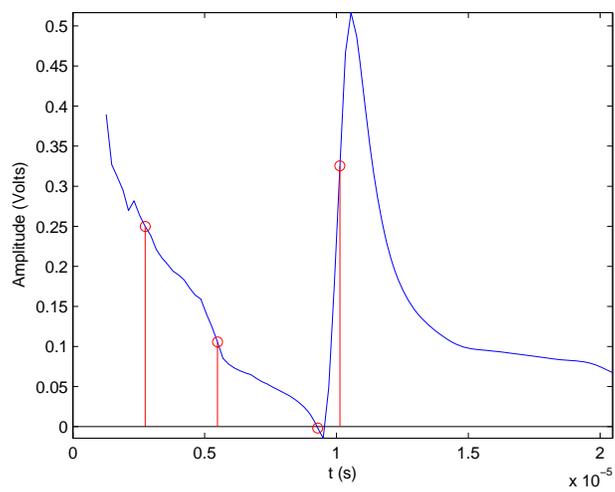


Figura 5.1: Saída do WFE para três valores de k_s : 1, 1.5 e 2.

5.1.2 IRA - *Impulse Response Analysis*

A seguir o método IRA será descrito através de exemplos. Com IRA, para enlaces seriais, é possível obter informação do comprimento total TL . Já para enlaces com *bridged-taps*, além do comprimento total, é possível determinar o número de *bridged-taps* e seus comprimentos. Nos dois exemplos mostrados a seguir nos quais é descrito o funcionamento de IRA, são usados dois enlaces com duas *bridged-taps*, pois a estimação do comprimento total é igual para enlaces seriais e enlaces com *bridged-taps*, mas a estimação das outras informações não é trivial. No primeiro caso, as duas *bridged-taps* tem comprimentos diferentes, e assim seus picos podem ser diferenciados facilmente, já que os picos referentes as duas *bridged-taps* estarão em posição diferentes, por percorrem distâncias diferentes. No segundo caso, as duas *bridged-taps* tem o mesmo comprimento e devido as reflexões causadas por essas *bridged-taps* percorrerem os mesmos comprimentos, apenas um pico é visível para ambas as *bridged-taps* na resposta impulsiva. Assim, nesse caso, é necessário um processo adicional para detectar as duas *bridged-taps*.

- **Primeiro exemplo – Duas *bridged-taps* com comprimentos diferentes**

Passo 1 - Estimação do comprimento total do enlace:

Na Figura 5.2 é mostrada a resposta impulsiva de um enlace, com as bordas de subida e descida encontradas pelo WFE. Devido a característica de propagação da resposta impulsiva, na qual o sinal no caminho direto do enlace chega primeiro, o primeiro pico é sempre referente ao comprimento total (TL) do enlace, que é obtido da borda de subida desse pico, através da equação:

$$TL(t_{init}) = v_c * t_{init}, \quad (5.2)$$

onde v_c é velocidade de propagação no cobre, que é 68% da velocidade da luz, e t_{init} é o tempo associado a borda de subida do primeiro pico presente na resposta impulsiva.

Passo 2 - Estimação do número de *bridged-taps* e seus comprimentos:

Os picos após o pico referente ao comprimento total do enlace podem estar relacionados a *bridged-taps* ou a espúrios devido a múltiplas reflexões. Para diferenciá-los é usado o seguinte algoritmo. Para os picos cujas bordas de subida e descida foram detectadas, é encontrado o vale presente logo antes da borda de subida desse pico. Na Figura 5.3 é mostrada a resposta impulsiva com a marcação em diferentes cores dos picos que podem estar relacionados a *bridged-taps*. Apesar não ter sido detectada a borda de descida do pico 3, como mostrado na Figura 5.3, fazendo que esse pico não seja considerado no restante da análise, ele foi destacado aqui apenas para ilustrar que mesmo que ele fosse

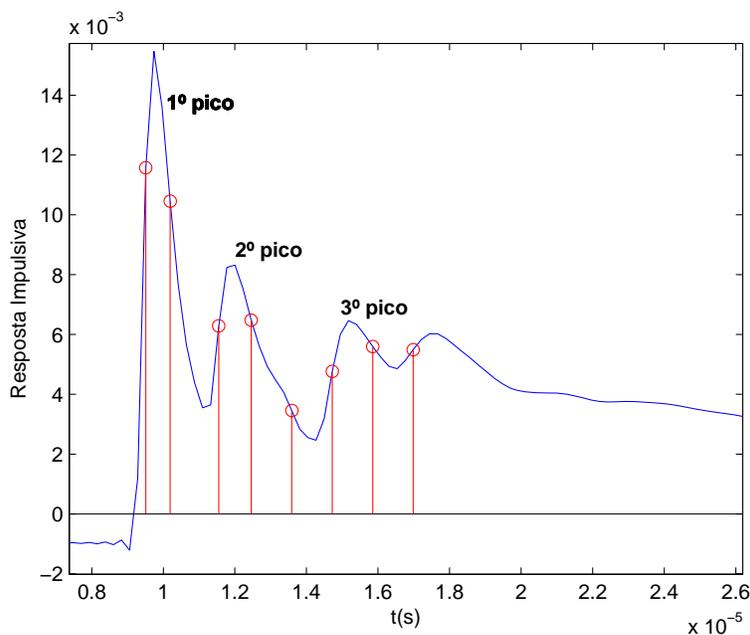


Figura 5.2: Resposta impulsiva com as bordas de seus picos identificados em vermelho.

considerado, o restante do algoritmo o descartaria como um pico espúrio. Como cada *bridged-tap* causa um vale proeminente seguido de um pico, o algoritmo de detecção de *bridged-tap* leva em conta a profundidade desse vale em relação ao seu pico. Quanto maior é essa profundidade maior a chance daquele pico ser referente a uma *bridged-tap*, já que para reflexões espúrias os vales praticamente inexistem. Os vales associados aos picos na Figura 5.3 foram destacados com a mesma cor que seu respectivo pico em linha tracejada. Para cada um dos picos, é calculada a diferença percentual entre a amplitude do pico e a amplitude de seu respectivo vale com a relação à amplitude do pico. Se essa diferença percentual for maior que 30%, esse pico é considerado relacionado a uma *bridged-tap*. Esse limiar foi determinado empiricamente através de testes. No caso em questão os valores percentuais foram:

- Preto: 37%;
- Laranja: 33%;
- Vermelho escuro: 10%.

Assim, apenas o pico número 3 não está relacionado a uma *bridged-tap*, sendo considerado um pico espúrio. Através da borda de subida do primeiro e segundo pico, é calculado o comprimento de ambas as *bridged-taps*.

Passo 3 - Estimação das bitolas do enlace:

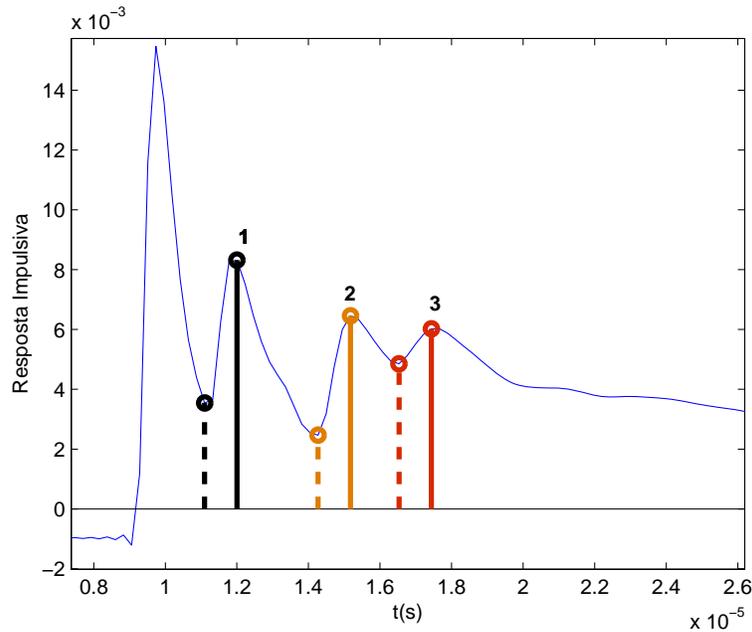


Figura 5.3: Os picos possivelmente relacionados a *bridged-taps*.

Das informações obtidas do comprimento total TL do enlace e das *bridged-taps*, pode ser obtido o seguinte enlace:

type = {serial, bridged-tap, serial, bridged-tap}

length = $[\frac{TL}{2}, L_{BT_1}, \frac{TL}{2}, L_{BT_2}]$

Deve-se notar que as *bridged-taps* são colocadas simetricamente ao longo do enlace, $\frac{TL}{2}$, já que o IRA não obtém informação sobre a posição das *bridged-taps*. L_{BT_1} e L_{BT_2} são os comprimentos das *bridged-taps* relacionados aos picos 1 e 2, respectivamente, que foram obtidos no **Passo 2**. Usando o enlace mostrado acima e um conjunto de sequência de bitolas, como por exemplo $[0,4, 0,4, 0,4, 0,4]$, $[0,4, 0,4, 0,4, 0,5]$, $[0,4, 0,5, 0,4, 0,4]$, e através de um modelo de linha é obtida a resposta impulsiva de cada um dos enlaces com essas sequências de bitolas. As sequências de bitolas testadas são combinações das bitolas padrões $[0,32 \ 0,40 \ 0,50 \ 0,63 \ 0,90]$ mm conforme o número de seções do enlace e a possibilidade de uma bitola ocorrer em dada posição no enlace (por exemplo, a bitola 0,32mm ocorre apenas próximo ao CO). A partir das respostas impulsivas obtidas para aquelas sequência de bitolas, é calculada a diferença percentual do pico relacionado ao comprimento total e aos picos das *bridged-taps* com relação aos respectivos picos da resposta impulsiva medida. Assim, $E_{TL}(\%)$ é o erro percentual do pico referente ao comprimento total com relação aos dados medidos, e $E_{BT_k}(\%)$ é o erro do k -ésimo pico relacionado a uma *bridged-tap*. A sequência de bitolas que possuir a menor média

dos erros dos picos, é a sequência escolhida para o enlace. Por exemplo, na Figura 5.4 é mostrada a comparação das curvas das três sequência de bitolas utilizadas como exemplo e a resposta impulsiva medida. Na Tabela 5.1 são mostrados os erros relacionados aos três picos. Como se pode ver a sequência com menor média dos erros dos picos foi [0.4 0.4 0.4 0.4]mm, que é escolhida para constituir o enlace mostrado anteriormente. É importante notar que os erros para essa sequência de bitola foram sempre menor que 5%. Se apresentasse valores maiores poderia indicar a existência de alguma *bridged-tap* adicional de mesmo comprimento que uma das *bridged-taps* encontradas. Nesse caso, um procedimento adicional teria de ser aplicado para fazer avaliação da existência dessa *bridged-tap*, como será explicado no exemplo 2.

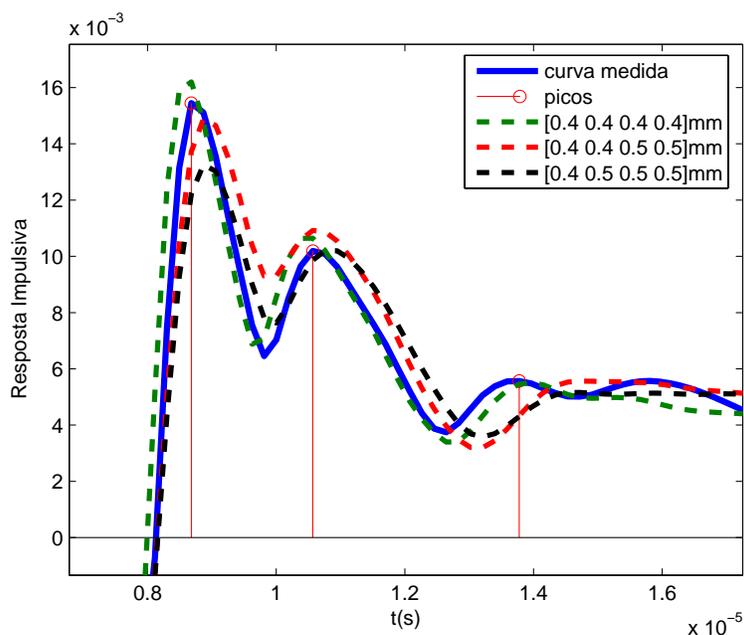


Figura 5.4: Determinação das bitolas das seções do enlace.

Tabela 5.1: Três sequência de bitolas e desvios relacionados a cada pico analisado.

Sequência de Bitolas	$E_{TL}(\%)$	$E_{BT_1}(\%)$	$E_{BT_2}(\%)$	Média (%)
[0.4, 0.4, 0.4, 0.4]	4.87	4.48	1.11	3.49
[0.4, 0.4, 0.4, 0.5]	1.88	2.10	10.3	4.76
[0.4, 0.5, 0.4, 0.4]	1.34	10.6	2.28	4.74

Na Tabela 5.2 é mostrado o enlace sob teste e o enlace estimado pelo IRA. O comprimento total estimado foi de 1.823km, erro de 1.464%. Deve-se notar que esse valor e os outros

parâmetros do enlace estimado são informações *a priori* que são melhoradas usando o método TIMEC descrito anteriormente.

Tabela 5.2: Resultados obtidos por IRA para o primeiro exemplo. *S* indica uma seção serial e *BT* indica uma *bridged-tap*.

TOPOLOGIA REAL			ESTIMAÇÃO DO IRA		
Tipo	g	<i>l</i>	Tipo	g	<i>l</i>
	(mm)	(km)		(mm)	(km)
S	0.4	0.75	S	0.4	0.912
BT	0.4	0.20	BT	0.4	0.198
S	0.4	1.10	S	0.4	0.912
BT	0.4	0.50	BT	0.4	0.535

- **Segundo exemplo – Duas *bridged-taps* com comprimentos diferentes**

Passo 1 - Estimação do comprimento total do enlace:

Na Figura 5.5 é mostrada a próxima resposta impulsiva que será analisada para exemplificar o funcionamento de IRA. As bordas de subida encontradas pelo WFE estão destacadas em vermelho. Da borda de subida do primeiro pico é encontrado o comprimento total TL do enlace.

Passo 2 - Estimação do número de *bridged-taps* e seus comprimentos:

Nesse caso, foram encontrados dois pares de bordas de subida e descida. É claro, a partir da Figura 5.5, que a borda de subida e descida do último par não estão relacionadas, ou seja, não fazem parte do mesmo pico. Como os picos e vales gerados ao longo da resposta impulsiva tem praticamente a mesma largura, a largura do pulso relacionado ao comprimento total é usada para verificar se as bordas pertencem a um mesmo pulso. Se a largura determinada por uma borda de subida e uma borda de descida for maior que 1,5 da largura do pulso relacionado do comprimento total, esse par é considerado não relacionado e assim as bordas pertencem a picos diferentes, sendo ignoradas pelo restante do algoritmo. Assim, na Figura 5.5 apenas um pico é detectado além do pico relacionado ao comprimento total. A diferença percentual desse pico e do vale correspondente é maior que 15%, determinando que esse pico é referente a uma *bridged-tap*. Através da borda de subida do pico, o comprimento da *bridged-tap* é determinado.

Passo 3 - Estimação das bitolas do enlace:

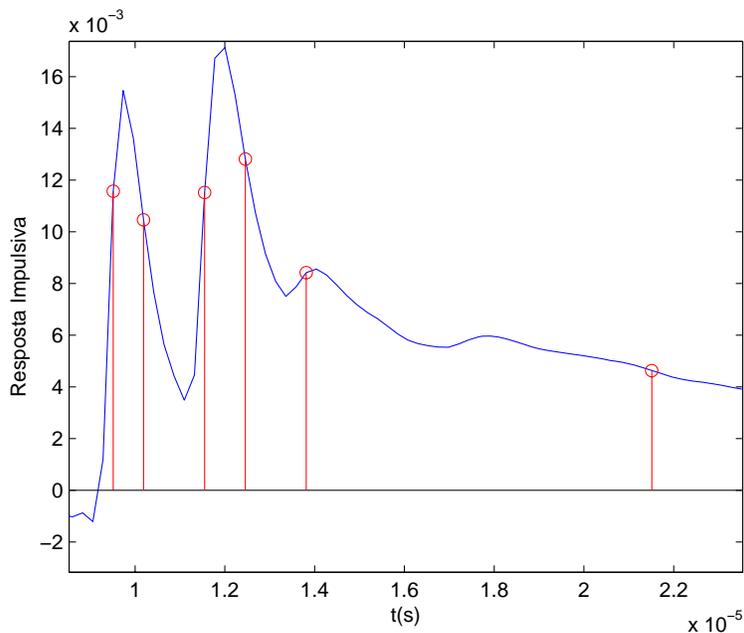


Figura 5.5: Picos possivelmente relacionados a *bridged-taps*.

Como apenas uma *bridged-tap* foi encontrada, a versão parcial do enlace está mostrado abaixo, onde L_{BT} é o comprimento da *bridged-tap*.

type = {serial, bridged-tap, serial }

length = $[\frac{TL}{2}, L_{BT}, \frac{TL}{2}]$

Seguindo o mesmo procedimento descrito para o exemplo 1, é determinado as bitolas do enlace, que são [0.4 0.4 0.4]mm. Na Figura 5.6 é mostrada a resposta impulsiva do enlace obtido comparada a curva medida. Os erros foram de 46.91% e 15.43% para os picos relacionados ao comprimento total e ao pico relacionado a *bridged-tap*, respectivamente. A média dos erros foi de 31.17%. Como um dos erros é maior que 15%, o algoritmo verificará a possibilidade da existência de outra *bridged-tap*.

Passo 4 - Verificar se existem outras *bridged-taps*:

Para testar a existência de uma *bridged-tap* extra, o seguinte enlace é constituído:

type = {serial, bridged-tap, serial, bridged-tap}

length = $[\frac{TL}{2}, L_{BT}, \frac{TL}{2}, L_{BT}]$

gauges = [0.4, 0.5, 0.4, g_x]

A *bridged-tap* adicional tem o mesmo comprimento que a primeira *bridged-tap*. A sua bitola, representada por g_x , é uma das bitolas que podem acontecer nessa posição do

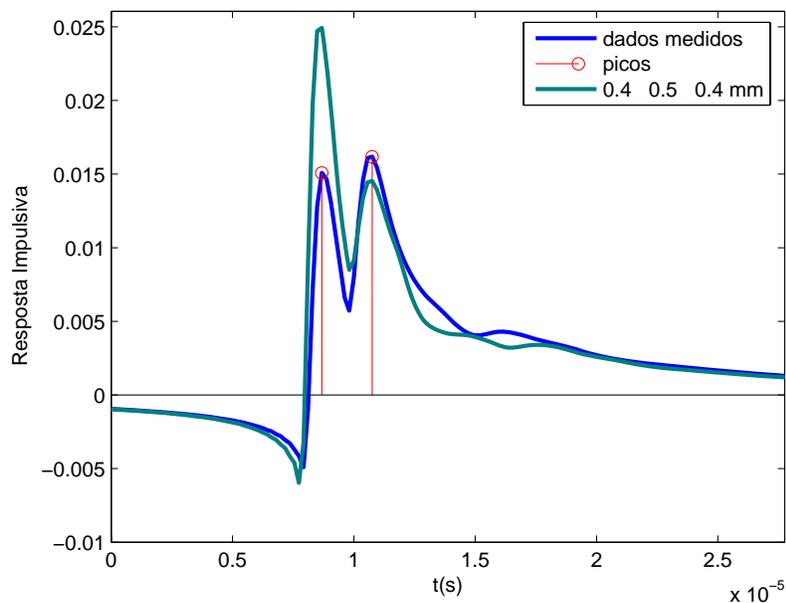


Figura 5.6: Comparação da curva medida e a obtida do enlace com as bitolas [0.4 0.5 0.4]mm.

enlace e que apresentam menor média dos erros dos picos. A primeira parte dessa sentença é para evitar que sequências de bitolas muito improváveis na prática ocorram, como por exemplo, 0.32mm próximo ao usuário. Para o enlace mostrado acima, o valor de g_x encontrado foi de 0,4mm. A resposta impulsiva do enlace, constituído pelas bitolas [0.4,0.5,0.4,0.4]mm, é mostrado na Figura 5.7. Pode-se notar que houve uma redução significativa dos erros associados aos picos do comprimento total e da *bridged-tap* em relação aos dados medidos, com valores de 6.73% e 10.03%, respectivamente. Como esses erros foram reduzidos a 20% dos erros quando a *bridged-tap* adicional não estava incluída, a existência dessa *bridged-tap* no enlace é muito provável. Conforme a variação do erro, a existência da *bridged-tap* adicional é avaliada como segue, baseada em testes empíricos:

- Se ambos os erros associados aos picos forem reduzidos, a probabilidade é considerada de pelo menos 85%, mas esse valor pode ser maior, se:
 - * Se ambos os erros caírem para 20% dos erros anteriores (para o enlace sem a *bridged-tap* extra), a probabilidade é de 99%;
 - * Se um dos erros cair abaixo de 20% do valor anterior, a probabilidade é de 95%;
 - * Se ambos os erros caírem em 50%, a probabilidade é de 90%;
- Se um dos erros diminuir e outro aumentar, mas esse aumento for menor que 50%,

- a probabilidade é:
- * 85%, se o erro diminuir para 20% do erro anterior correspondente;
 - * 80%, se o erro cair abaixo de 50% do valor anterior;
- Se um dos erros diminuir e outro aumentar, e esse aumento for maior que 50%, a probabilidade é:
- * 70%, se o erro diminuir para 20% do erro anterior correspondente;
 - * 50%, se um dos erros cair abaixo de 40% do valor anterior;
 - * 30%, se um dos erros cair abaixo de 50% do valor anterior;

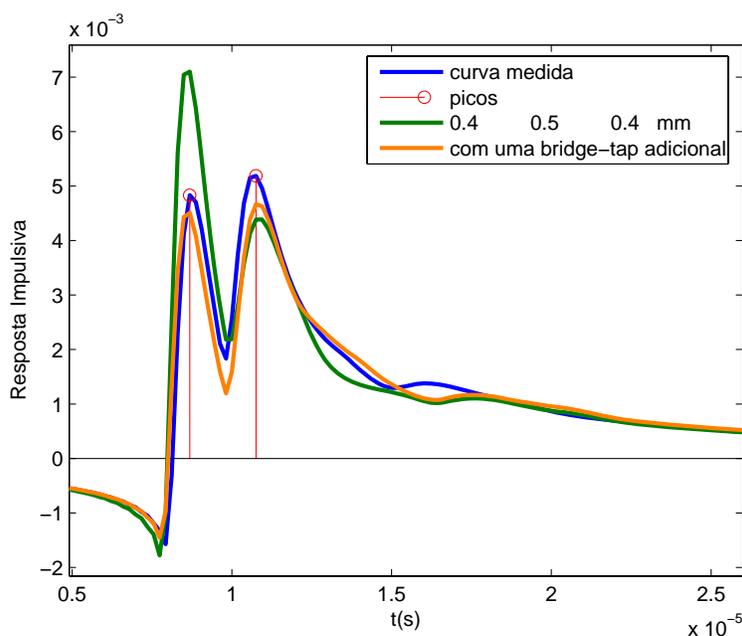


Figura 5.7: Comparação da curva medida e a obtida do enlace com a *bridged-tap* adicional.

Para qualquer outra alternativa não considerada, a probabilidade é menor que 30%. Na Tabela 5.3 é mostrado o enlace de teste e o enlace estimado pelo IRA. Novamente, deve-se destacar que esse resultado é apenas uma estimativa inicial a ser passada para o método principal que é o TIMEC.

5.1.3 First Section Finder (FSF)

Como pode ser visto da técnica descrita em [11], chamada aqui SELT-tdr, medidas de TDR podem ser utilizadas para identificação de topologias de enlaces. Apesar da técnica SELT-tdr ser simples de implementar, ela tem uma severa limitação na identificação de enlaces

Tabela 5.3: Resultados obtidos aplicando IRA para dados simulados. S indica uma seção serial e BT indica uma *bridged-tap*.

TOPOLOGIA REAL			ESTIMAÇÃO DO IRA		
Tipo	g	l	Tipo	g	l
	(mm)	(km)		(mm)	(km)
S	0.4	0.75	S	0.4	0.911
BT	0.4	0.20	BT	0.5	0.218
S	0.4	1.10	S	0.4	0.911
BT	0.4	0.20	BT	0.4	0.218

seriais com mais de uma seção, como foi demonstrado em [61], pois devido a erros cumulativos durante o processo de *de-embedding* (desconto da resposta de uma seção identificada no reflectograma TDR), as estimativas das seções restantes são afetadas, assim como a estimativa do comprimento total. No entanto, para identificação da topologia da primeira seção do enlace, SELT-tdr é bastante precisa. Além disso, enquanto uma das principais deficiências de SELT-tdr é determinar o início de um eco com precisão, que pode impedir o desempenho desse método devido a erros no processo de *de-embedding*, o WFE obtêm essas características do sinal mais facilmente e com mais precisão.

Como foi visto na seção 5.1.1, a determinação do fator de sensibilidade k_s do WFE é baseado na identificação das bordas de subida e descida do pico referente a reflexão no final do enlace. Mas ao mesmo tempo que se determina k_s a ser usado pelo WFE na determinação dos ecos, obtêm-se o valor do comprimento total do enlace TL , da seguinte equação:

$$TL(t_{init}) = (v_c * t_{init})/2, \quad (5.3)$$

onde v_c é velocidade de propagação no cobre, que é 68% da velocidade da luz, e t_{init} que é tempo de chegada do eco, que para o comprimento total TL é o tempo associado a borda de subida do pico referente ao final do enlace.

Outra informação obtida é sobre a existência de *bridged-taps*. Se houver vales com amplitudes negativas na resposta impulsiva do enlace, isto indica a existência de *bridged-taps* no enlace. Para cada vale de amplitude negativa existe um *bridged-tap* associada. Assim n_{bt} *bridged-taps* vão ser determinados por n_{bt} vales de amplitude negativa. Da borda de descida do vale é obtida informação sobre a posição da *bridged-tap* ao longo do enlace, indicada aqui por **posbt**. Em alguns casos, associado ao vale tem-se um pico de amplitude positiva, do

qual se pode obter informação sobre o comprimento da *bridged-tap* $\overline{l_{bt}}$. Uma informação que pode ser obtida tanto para enlaces com *bridged-tap* quanto para enlaces seriais, além do comprimento total, é a topologia da primeira seção do enlace. Nesse caso, a estimação da bitola é realizada como em []. Essa é a única parte da técnica SELT-tdr usada por FSF, mas ainda sim modificada e mais simples. Para encontrar a bitola da primeira seção o erro médio quadrático (em inglês mse - *mean square error*) é usado para comparar a resposta TDR medida com as respostas TDR de enlaces com uma das bitolas disponíveis e de comprimentos infinitos. No caso de serem usadas as bitolas [0,32, 0,40, 0,50, 0,63, 0,90]mm, a resposta TDR de cinco enlaces de comprimento infinito serão comparadas a resposta TDR medida. Na Figura 5.8 é mostrada a resposta TDR medida e as respostas TDR para cada uma das bitolas diferentes. Essa comparação é realizada até um tempo t_{final} , que indica o final do eco referente a primeira seção. A bitola que tiver menor valor do mse, é escolhida como a bitola da primeira seção do enlace. No caso de SELT-tdr é necessário usar artifícios relativamente complexos para a aplicação do mse, já que não é possível determinar o valor de t_{final} sem antes aplicar o processo de *de-embedding*. Por outro lado, usando WFE o início de cada eco do reflectograma TDR são todos determinados ao mesmo tempo, e dessa forma o valor de t_{final} do eco da primeira seção é facilmente determinado. Portanto, o método para encontrar a bitola da primeira seção torna-se bem mais simples do que o usado em SELT-tdr.

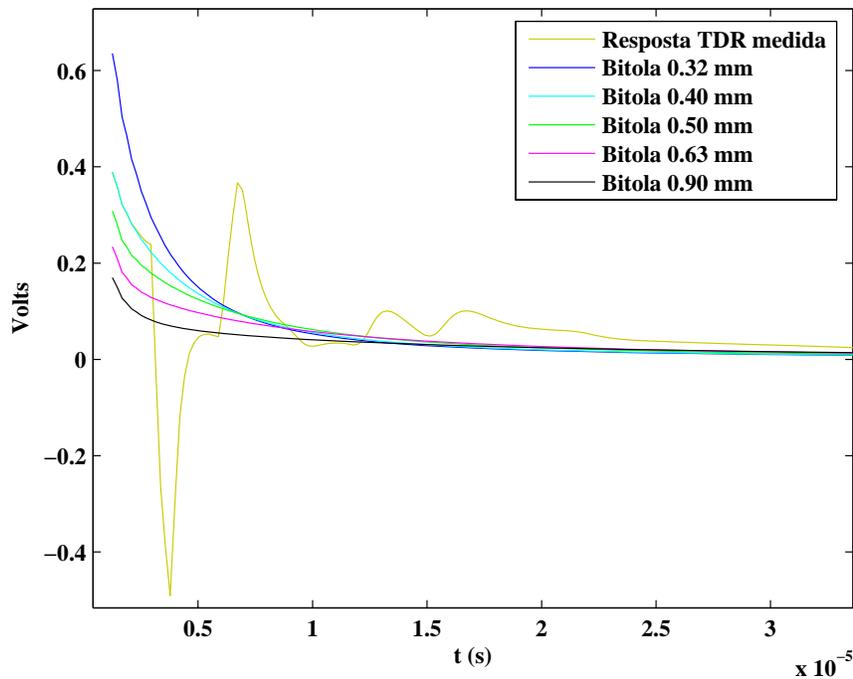


Figura 5.8: Método para determinar a bitola da primeira seção de um enlace.

O comprimento da primeira seção \bar{l}_1 é obtido de pontos de borda. Para os enlaces utilizados no experimento, testes mostraram que se apenas as bordas de subida e descida do pico referente ao final do enlace forem encontradas, esse enlace tem apenas uma seção. Assim, para esse caso, com a bitola encontrada como mostrado anteriormente e com o cálculo do comprimento total, esse enlace está totalmente determinado. Esse caso é mostrado na Figura 5.9, onde a resposta TDR do enlace de bitola 0,5mm e com comprimento de 500m foi utilizado no teste. Se outras bordas de subida de amplitude positiva (já que valores negativos indicam *bridged-taps* e são tratados separadamente) forem encontradas além do pico de final de enlace, provavelmente o enlace possui mais de uma seção. Não se pode ter certeza pois bordas espúrias podem persistir mesmo para valores baixos de sensibilidade do WFE, indicando seções extras que na verdade não existem. Esses espúrios podem também aumentar o erro na estimação do comprimento da primeira seção. Mas como no caso do IRA, essas informações são estimativas iniciais que podem ser melhoradas pelo TIMEC, e que mesmo com erros, podem significar um atalho na pesquisa do enlace correto. De qualquer forma, mantendo um grau de liberdade adequado do TIMEC, informações incorretas obtidas pelo FSF (ou IRA) podem ser facilmente descartadas e as informações corretas apresentando valiosa contribuição para o desempenho de TIMEC.

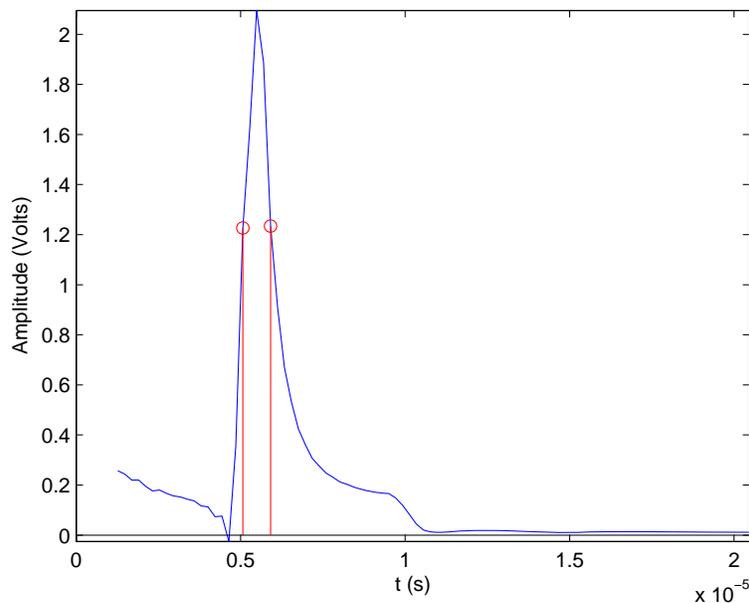


Figura 5.9: Resposta impulsiva para o enlace de 500m e bitola 0.5mm.

Se apenas uma borda positiva for encontrada antes da borda de subida do pico de final de enlace, como mostrado na Figura 5.10, o comprimento da primeira seção é determinada por essa borda através da equação (5.3), onde t_{init} é o tempo associado a ela. Se a posição

da *bridged-tap* for maior que o comprimento da primeira seção, isto significa que pode existir até duas seções antes da *bridged-tap*. Essa informação também é usada pelo IRA. No caso da Figura 5.10 o valor encontrado para o comprimento da primeira seção foi de 552m, erro de 10,4% na estimação desse parâmetro.

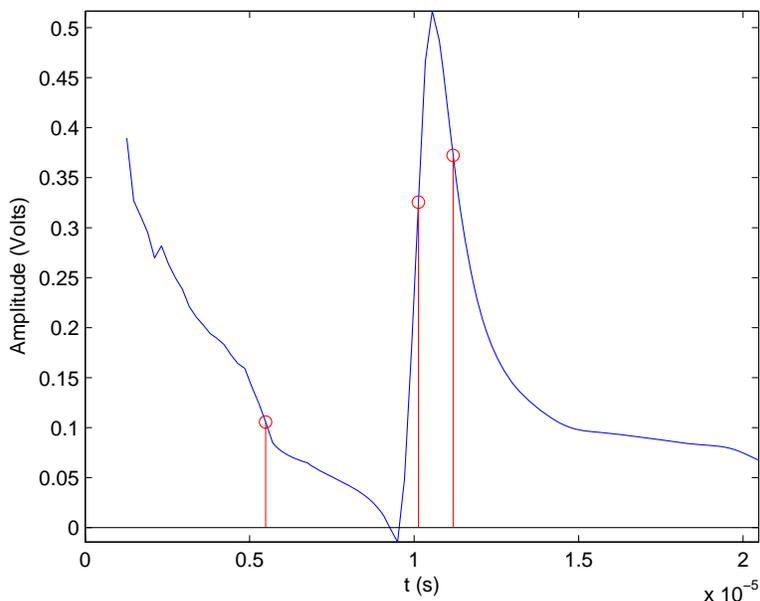


Figura 5.10: Resposta impulsiva com três pontos identificados.

Outro caso que pode ocorrer é o mostrado na Figura 5.11, onde dois pontos de borda foram encontrados antes do pico de final de enlace. Uma dessas bordas é um espúrio que é identificado pelo WFE incorretamente como um eco verdadeiro. Como não é possível determinar qual é a borda espúria, ambos os comprimentos são considerados, sendo chamados de \bar{l}_1 e \bar{l}_2 , o primeiro referente ao comprimento de menor valor e o segundo ao comprimento de maior valor. Ambos valores são passados ao TIMEC, que realiza a pesquisa entre $[\bar{l}_1 - \Delta\%, \bar{l}_2 + \Delta\%]$, onde $\Delta\%$ é um valor percentual de até 10%. Dessa forma, o TIMEC ainda vai considerar valores que podem está distantes do valor real do comprimento da seção (algumas vezes o dobro do valor real), mas ainda sim essa faixa de pesquisa é sempre bem menor que a faixa utilizada normalmente para pesquisar o comprimento de uma seção (que é de 200m a 5.5km nas simulações realizadas no capítulo de resultados). Com isso o tempo necessário para estimar o comprimento dessa seção é bastante reduzido.

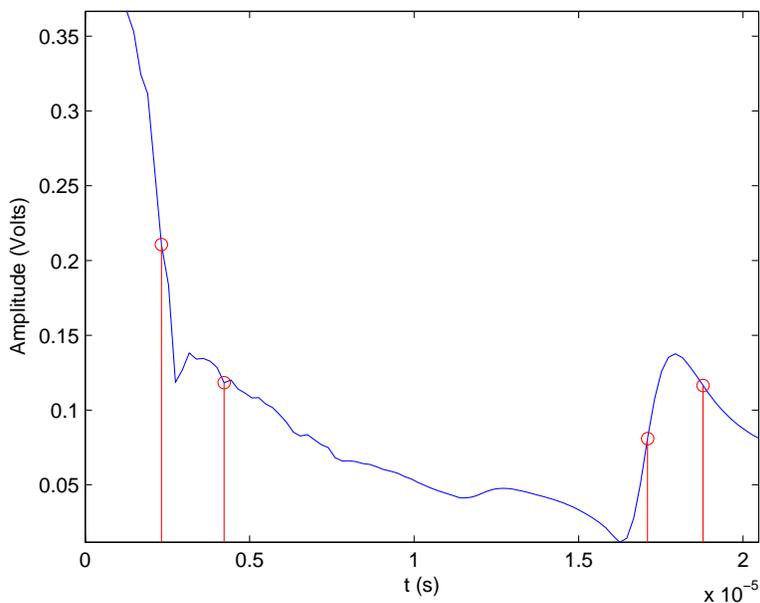


Figura 5.11: Resposta impulsiva com 4 pontos identificados.

5.2 Conclusões do Capítulo

Nesse capítulo foram descritas as ferramentas de pré-processamento responsáveis por extrair informação *a priori* sobre o enlace. A primeira trata-se de uma técnica que realiza a análise da resposta impulsiva, chamada IRA, e que pode determinar informações como o comprimento total e o número de *bridged-taps*. Essa técnica é bastante simples e tem tempo de processamento bastante reduzido. A outra técnica descrita foi o FSF que usa medidas de TDR para obter informações sobre o enlace. Como o IRA, essa técnica também é simples de implementar e tem custo computacional pequeno. Através do FSF, podem-se obter informações sobre a topologia da primeira seção e informações sobre as *bridged-taps* existentes no enlace. Através das informações dessas duas técnicas, o TIMEC pode encontrar mais rapidamente a topologia do enlace e com uma precisão bem maior. No próximo capítulo serão apresentados os resultados obtidos, primeiro de cada módulo trabalhando separadamente e a seguir os resultados de todos os módulos atuando em conjunto.

Capítulo 6

Resultados

Nesta seção serão apresentados os resultados obtidos a partir da metodologia de identificação de enlaces LQ-SUITE. A primeira seção mostra a comparação entre o LQ-SUITE e um método de identificação baseado em reflectometria no domínio do tempo. Nas seções seguintes, serão descritos os resultados de cada um dos módulos do LQ-SUITE individualmente e por fim do próprio LQ-SUITE. Na última seção, são apresentados os comentários finais do capítulo.

6.1 Comparação de TIMEC e *SELT-tdr*

Esta seção mostra a comparação dos resultados obtidos pelo método TIMEC e a técnica baseada em TDR chamada *SELT-tdr* [11], que são descritos em [61], mais com o acréscimo de outros enlaces. Os módulos IRA e FSF ainda estavam em desenvolvimento, e assim apenas TIMEC podia ser usado. A versão de TIMEC usada nesse artigo era bastante restrita, não possuindo a maioria das melhorias descritas no Capítulo 4, como a técnica de otimização local, o processo de realimentação de seções e o processo de decodificação *typegauge*. Dessa forma será possível também avaliar o impacto da aplicação dessas melhorias e também da utilização de informações *a priori* obtidas por IRA e FSF. Nas simulações foram utilizados dados simulados obtidos de modelos de linha, ou seja, ao invés de dados medidos da função de transferência \hat{H} e do parâmetro de espalhamento \hat{S}_{11} para um dado enlace de teste, são utilizados dados gerados pela aplicação das Equações 2.10 e 2.11, juntamente com a teoria ABCD padrão e modelos de linha. Em outras palavras, a partir dos valores verdadeiros dos parâmetros do enlace Θ^\dagger são geradas as curvas alvos $\hat{H} = H^\dagger$ e $\hat{S}_{11} = S_{11}^\dagger$. Sob estas condições controladas, sem o desvio das curvas alvos e o modelo, e também sem a influência de fontes de ruído, é possível encontrar a estimativa exata dos parâmetros do enlace $\Theta^* = \Theta^\dagger$.

Nas simulações realizadas, os enlaces de teste foram escolhidos entre os enlaces padrões definidos em [58]: ETSI #1 ao ETSI #8 mostrados no Apêndice B, e também um enlace teste definido em [11] que apresenta três seções de bitola 0,5 mm, sendo a primeira uma seção serial de 300 metros, a segunda uma *bridged-tap* de comprimento 300 metros e a última uma seção serial de 900 metros. Esse último enlace será chamado SBTS. Para enlaces de teste apresentando as bitolas sempre crescentes, que incluem os enlaces ETSI com exceção do ETSI #8, a técnica de decodificação de bitolas *gaugesort* foi usada; para os outros enlaces, foi usada a técnica *freegauge*. Essa era uma limitação inicial do método proposto, porque era necessário saber que não haveria repetição de uma mesma bitola ao longo do enlace e que elas sempre seriam crescentes a partir do CO para aplicar a técnica *gaugesort*, que restringe de forma mais inteligente o espaço de busca e assim melhora a eficiência do método. Essa limitação foi superada pela aplicação da técnica *typegauge*, que não necessita dessas informações. A técnica *typegauge* foi utilizada nas simulações descritas nas seções mais a frente.

O TIMEC foi simulado dez vezes para cada enlace de teste de forma a avaliar o seu intervalo de confiança e a sua precisão. Por ser um método baseado em uma técnica estocástica de otimização, é importante avaliar se o TIMEC é robusto e conduz aproximadamente aos mesmos resultados para diferentes execuções independentes. Portanto, ao invés de simplesmente ser dada uma estimativa do comprimento l do enlace, os resultados obtidos por TIMEC serão dados em termos da média e do desvio padrão considerando todas as suas execuções. Será também avaliado o *comprimento total* do enlace, que é a soma do comprimento de todas as seções seriais do enlace.

Para todas as simulações, foram usados $R = 30$ indivíduos e foi utilizado o cruzamento aritmético em conjunto com a mutação gaussiana, e também a seleção por torneio com $t = 3$ (número de indivíduos por competição). O AG empregado usa o algoritmo de controle auto-adaptativo dos parâmetros descrito na Seção 4.2.4. Os valores máximos para a probabilidade de cruzamento P_c e para a probabilidade mutação P_m foram configurados para 0,85 e o desvio padrão máximo Σ_m foi configurado para 0,75. Adicionalmente, $\Delta N = 50$, w é igual a 0,05, δ_{p_c} e δ_{p_m} são configurados para 0,005, e δ_{σ_m} é configurado para 0,01. As faixas de valores dos parâmetros otimizados são descritos na Tabela 6.1. As bitolas consideradas nessas simulações foram as bitolas usadas no padrão europeu [0,32, 0,40, 0,50, 0,63, 0,90] mm.

O método *SELT-tdr* foi escolhido por ser recente na literatura e apresentar resultados melhores que outras técnicas de identificação de enlaces, além de ser de fácil implementação. Deve ser notado que a técnica *SELT-tdr* foi implementada estritamente seguindo o processo descrito em [11], com exceção da técnica de detecção de singularidades do reflectograma, que não é descrita nesse artigo. Para qual, outra técnica, baseada em derivadas, foi desenvolvida para estimar o tempo de chegada dos ecos verdadeiros. Portanto, para a validação da imple-

Tabela 6.1: Faixa de variação dos parâmetros sob otimização.

Parâmetro	mínimo	máximo
Comprimento da seção serial (km)	$L_{\min} = 0,02$	$L_{\max} = 7$
Comprimento da seção <i>bridged-tap</i> (km)	$L_{\min} = 0,05$	$L_{\max} = 0,8$
Número de seções	$N_s^{\min} = 1$	$N_s = 5$
Número de <i>bridged-taps</i>	$N_b^{\min} = 0$	$N_b = 2$

mentação da técnica *SELT-tdr* realizada para esta proposta, alguns testes com enlaces usados em [11] e [14] foram conduzidos e resultados similares foram encontrados aos descritos nessas referências. De forma a ilustrar esse estágio de validação, o resultado para o enlace arbitrário SBTS, que serve de exemplo de aplicação da técnica *SELT-tdr* em [11], é apresentado na Tabela 6.2. Além disso, foram usados os enlaces do padrão europeu: ETSI # 1 ao ETSI # 8. Os resultados das simulações descritos na Tabela 6.2, mostram o comprimento l , a bitola g e o tipo de cada seção dos enlaces estimados pelo método proposto TIMEC e pela técnica *SELT-tdr*, além dos enlaces testes. O erro percentual $\Delta l(\%)$ do comprimento das seções dos enlaces estimados pela técnica *SELT-tdr* é mostrado na Tabela 6.2. Para o método TIMEC, são calculados o erro médio percentual $\overline{\Delta l(\%)}$, o comprimento médio \bar{l} e o desvio padrão σ para cada seção do enlace, considerando as dez execuções independentes do método.

Os resultados obtidos por método *SELT-tdr* demarcam duas características inerentes a esse método: ele é bem adaptado para detectar *bridged-taps* mas não é bem sucedido em detectar mudanças de bitolas. Uma mudança de bitola constitui uma mudança de impedância que é frequentemente pequena em comparação, por exemplo, a mudanças de impedância devido a uma *bridged-tap*. Algumas mudanças de bitola tem um coeficiente de reflexão mais proeminente que outras, como mostrado na Figura 6.1. Nessa figura, a magnitude do coeficiente de reflexão para todas as mudanças de bitolas dos enlaces de teste considerados é mostrada como função da frequência. Pode-se ver a partir figura que especialmente as mudanças de bitola 0.32/0.4 mm, 0.32/0.5 mm e 0.5/0.9 mm apresentam valores elevados para o coeficiente de reflexão para uma ampla faixa de frequência. Consequentemente, essas mudanças de bitolas são mais facilmente detectadas. O enlace ETSI # 7 é constituído de duas dessas mudanças de bitolas e como se pode ver pela Tabela 6.2, esse enlace é corretamente identificado por *SELT-tdr*, exceto pela bitola da segunda seção. Também pode ser observado da Fig. 6.1 que as mudanças de bitola 0.5/0.63 mm e 0.63/0.9 mm tem o valor mais baixo do coeficiente de reflexão, causando perturbações pequenas na curva TDR. Isso se reflete nos resultados de

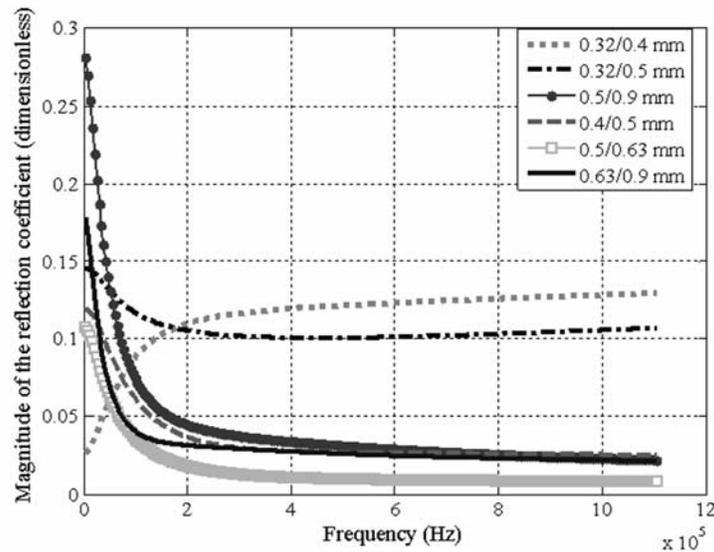


Figura 6.1: Magnitude dos coeficientes de reflexão como função da frequência relacionados a descasamentos de impedância a mudanças de bitolas encontradas nos enlaces ETSI.

SELT-tdr para os enlaces ESTI # 4, ESTI # 5 e ESTI # 6 onde somente uma ou duas seções seriais são detectadas, com um erro da estimação do comprimento total de 26 %, 13 % e 4 %, respectivamente. Para os enlaces de uma seção, isto é, enlaces ETSI # 1 e ETSI # 2, *SELT-tdr* apresentou um erro do comprimento total de 7 %, com suas bitolas corretamente encontradas.

Como explicado anteriormente, TIMEC usa duas funções custos, uma relacionada a medidas de uma porta e a outra a medidas de duas portas. Da Tabela 6.2 nota-se que o número seções para todos os enlaces de teste são corretamente estimadas com TIMEC, exceto para o enlace ETSI # 5, para o qual a última seção não foi estimada. Foi obtida a estimação perfeita do enlace ETSI # 8, com as *bridged-taps* encontradas em suas posições corretas, isto é, na segunda e quarta seção. Para o ETSI # 4, os comprimentos da terceira e quarta seções do enlace ficaram acima e abaixo dos valores corretos, respectivamente. O erro percentual médio do comprimento total para esse enlace foi de 2,0 %. Para o ETSI # 5, apesar da última seção não ser identificada, o erro do comprimento total foi de apenas 1,3 %, já que a segunda e terceira seções absorveram o comprimento da última seção perdida. Para o ETSI # 6, a bitola da última seção está incorretamente estimada e o erro percentual médio do comprimento total é de 1,24 %. Como pode ser notado pelos resultados apresentados, as estimações das seções mais distante do CO apresentam resultados de menor qualidade e as estimações das primeiras seções, próximas do CO, apresentam alta qualidade, com erro igual a zero ou próximo de zero. Isso se deve a influência do parâmetro de espalhamento S_{11} obtido a partir do CO,

que apresenta ecos mais intensos das reflexões das seções mais próximas do ponto de medição, facilitando a estimação dessas seções, mas que diminuem para seções mais afastadas, tornando mais difícil a estimação dessas seções. Já a função de transferência tem uma maior influência na estimação do comprimento total do enlace, pois esse parâmetro está fortemente relacionado a atenuação.

Tabela 6.2: Comparação entre *SELT-tdr* [11] e TIMEC usando duas funções objetivos baseadas em S_{11} e H . S indica um enlace serial, BT indica uma *bridged-tap* e TL trata-se do comprimento total.

Enlace de teste			<i>SELT-tdr</i>				TIMEC(S_{11}, H)				
Tipo	g (mm)	l (m)	Tipo	g (mm)	l (m)	Δl (%)	Tipo	g (mm)	l (m)	$\overline{\Delta l}$ (%)	σ (m)
S	0,40	2,55	S	0,4	2,707	6,16	S	0,40	2,55	0,0	0,0
S	0,50	3,40	S	0,5	3,617	6,37	S	0,50	3,40	0,0	0,0
S	0,40	1,30	S	0,4	2,664	105	S	0,40	1,30	0,0	0,0
S	0,50	1,25	–	–	–	–	S	0,50	1,401	12,24	0,11
S	0,63	0,50	–	–	–	–	S	0,90	0,387	23,0	0,11
S	0,32	0,20	S	0,32	0,217	8,35	S	0,32	0,20	0,0	0,0
S	0,40	0,90	S	0,40	2,535	181,69	S	0,40	0,90	0,0	0,0
S	0,50	1,50	S	–	–	–	S	0,50	1,563	4,20	0,08
S	0,63	0,50	S	–	–	–	S	0,63	404,0	19,20	0,11
S	0,40	1,45	S	0,40	2,362	62,89	S	0,40	1,45	0,0	0,0
S	0,50	0,75	–	–	–	–	S	0,50	0,925	23,60	0,09
S	0,63	0,50	–	–	–	–	S	0,63	0,769	55,00	0,14
S	0,90	0,50	–	–	–	–	–	–	–	–	–
S	0,40	1,30	S	0,40	2,664	104,90	S	0,40	1,3	0,0	0,0
S	0,50	1,25	–	–	–	–	S	0,50	1,401	12,24	0,11
S	0,63	0,50	–	–	–	–	S	0,90	0,387	23,00	0,11
S	0,32	0,20	S	0,32	0,217	8,35	S	0,32	0,20	0,0	0,0
S	0,50	0,60	S	0,40	0,672	11,95	S	0,50	0,60	0,0	0,0
S	0,90	4,00	S	0,90	3,10	22,54	S	0,90	4,0	0,0	0,0
S	0,40	0,75	S	0,40	0,737	1,77	S	0,40	0,75	0,0	0,0
BT	0,40	0,50	BT	0,40	0,108	78,34	BT	0,40	0,50	0,0	0,0
S	0,40	1,10	S	0,63	0,390	64,55	S	0,40	1,10	0,0	0,0
BT	0,40	0,50	BT	0,40	0,087	477,4	BT	0,40	0,50	0,0	0,0
S	0,50	0,30	S	0,50	0,303	1,07	S	0,50	0,3001	0,03	0,0
BT	0,50	0,30	BT	0,50	0,303	1,07	BT	0,50	0,3001	0,04	0,0
S	0,50	0,90	S	0,50	0,909	1,06	S	0,50	0,9002	0,02	0,0

6.2 Resultados do LQ-SUITE

Nas próximas seções serão apresentados os resultados de cada módulo do LQ-SUITE e por fim os resultados dos módulos trabalhando em conjunto. Quando utilizado o TIMEC, as configurações são as mesmas dadas na seção anterior. São usados dados simulados e dados medidos nos testes. O primeiro caso é usado para avaliar a técnica sob condições ideais e assim verificar qual o comportamento potencial da técnica. O segundo caso é para avaliar a técnica sob condições mais próximas das condições práticas, onde os modelos não se ajustam perfeitamente aos dados medidos.

6.2.1 Dados Simulados

Nas simulações realizadas para dados simulados, os enlaces de teste foram os enlaces padrões definidos em [58]: ETSI #1 ao ETSI #8 mostrados no Apêndice B. Dessa forma, os resultados obtidos podem ser comparados ao resultados apresentados na seção anterior, onde apenas o TIMEC foi utilizado.

6.2.1.1 Resultados do IRA

Na Tabela 6.3 são mostrados os resultados de IRA para os enlaces ETSI, usando dados simulados. Como previamente citado, quando um enlace é identificado como enlace serial, o algoritmo apenas retorna seu comprimento total. Nenhuma outra informação desses enlaces é possível obter na versão atual do IRA. Provavelmente, em versões mais avançadas de IRA, não será possível obter mais informações além dessa, pois como explicado para o caso de *SELT-tdr*, as mudanças de bitolas causam apenas leves impressões nas respostas de tempo analisadas, seja de TDR ou da resposta impulsiva, e que muitas vezes não podem ser diferenciadas de espúrios ou ruídos. Para esses enlaces, o erro máximo da estimação do comprimento total é de 2.27%. Para o enlace ETSI #8, que tem duas *bridged-taps*, IRA estimou corretamente o número de *bridged-taps* (probabilidade do enlace ter duas *bridged-taps* de 90%), com erro de 17.5% para estimação do comprimento dessas *bridged-taps*. Para esse enlace, apenas a bitola da primeira *bridged-tap* está incorretamente estimada (0,5 mm ao invés de 0,4 mm).

Esses resultados podem ser usados como informação *a priori* pelo TIMEC. No caso de enlaces seriais, é informado que o enlace teste é serial e o seu comprimento total. Já para o enlace ETSI #8, é informado que o enlace tem duas *bridged-taps* e seus comprimentos, além do comprimento total do enlace. Como explicado na Seção 4.2.2, qualquer valor de informação *a priori* é ainda otimizado por TIMEC, mas em torno do valor fornecido. Por exemplo, $\pm 20\%$

do valor do comprimento total fornecido por IRA. Para a bitola, que é um valor discreto, são considerados apenas os valores consecutivos a bitola encontrada por IRA, além dessa própria bitola. Assim, o valor da bitola da primeira *bridged-tap* do enlace pode ser corrigida por TIMEC. Além disso, TIMEC pode reposicionar as *bridged-taps* em suas posições corretas.

Tabela 6.3: Resultados obtidos aplicando IRA aos dados simulados dos enlaces ETSI. S indica um enlace serial, BT indica uma *bridged-tap* e TL trata-se do comprimento total.

ENLACE DE TESTE			ESTIMAÇÃO DO IRA					
Tipo	g (mm)	l (km)	Tipo	g (mm)	\bar{l} (km)	E_l (%)	TL (km)	erro TL (%)
S	0,40	2,55	Enlace Serial				2,58	1,18
S	0,50	3,40	Enlace Serial				3,49	2,71
S	0,40	1,40	Enlace Serial				2,93	1,03
S	0,50	1,50	Enlace Serial					
S	0,32	0,20	Enlace Serial					
S	0,40	0,90	Enlace Serial				3,17	2,32
S	0,50	1,50	Enlace Serial					
S	0,63	0,50	Enlace Serial					
S	0,40	1,45	Enlace Serial					
S	0,50	0,75	Enlace Serial				3,21	0,31
S	0,63	0,50	Enlace Serial					
S	0,90	0,50	Enlace Serial					
S	0,40	1,30	Enlace Serial					
S	0,50	1,25	Enlace Serial				3,09	1,31
S	0,63	0,50	Enlace Serial					
S	0,32	0,20	Enlace Serial					
S	0,40	0,60	Enlace Serial				4,8	0,10
S	0,90	4,0	Enlace Serial					
S	0,40	0,75	S	0,40	0,911	–		
BT	0,40	0,50	BT	0,50	0,535	7,0	1,822	1,51
S	0,40	1,10	S	0,40	0,911	–		
BT	0,40	0,50	BT	0,40	0,535	7,0		

6.2.1.2 Resultados do FSF

Na Tabela 6.4 é mostrado os resultados obtidos por FSF. No caso de dados simulados os picos de final de enlace não foram detectados, e assim o valor do comprimento total não foi encontrado, exceto para o enlace que possui *bridged-taps*, o ETSI#8. Além disso, o valor máximo do fator de sensibilidade do WFE K_s foi utilizado, já que todos os valores intermediários testados nenhum deles alcançou o critério definido: encontrar o pico de final de enlace, conforme método descrito na Seção 5.1.1. Exceto para os enlaces com bitolas 0,32mm, o erro da estimativa do comprimento da primeira seção foi abaixo 10%. Para os enlaces ETSI # 4 e ETSI # 7, o erro foi de 59%. Provavelmente o comprimento pequeno da seção e sua alta atenuação, causam múltiplas reflexões muito próximas umas das outras e do eco verdadeiro, dificultando a estimativa das bordas do eco real pelo WFE. Todas as bitolas da primeira seção foram corretamente encontradas. Além disso, o FSF identificou corretamente os enlaces seriais e aquele que possui *bridged-tap*. Para o enlace ETSI # 8, apenas uma *bridged-tap* foi identificada, mas não o seu comprimento. Como pode ser ver na Tabela 6.4, a *bridged-tap* se encontra logo após a primeira seção, já que o comprimento da primeira seção e a posição da *bridged-tap* possuem valores iguais. O erro desses comprimentos foi de 7.5%. Para ETSI # 8, o erro do comprimento total foi 27,8%, valor considerado elevado. No entanto, o comprimento total é obtido do IRA, que é especializado para obter esse parâmetro.

Deve-se notar que os valores obtidos na Tabela 6.4, são estimativas iniciais e todos esses valores, inclusive o valor das bitolas e o número de *bridged-taps* (para enlaces que possuem *bridged-taps*) serão otimizados por TIMEC.

Tabela 6.4: Resultados obtidos aplicando FSF aos dados simulados dos enlaces ETSI. O erro percentual é mostrado ao lado da estimativa. *S* indica um enlace serial, *BT* indica uma *bridged-tap* e *TL* trata-se do comprimento total.

ENLACE DE TESTE			RESULTADOS DO FSF						
Tipo	<i>g</i> (mm)	<i>l</i> (km)	nbt	posbt (km) (%)	\bar{l}_{bt} (km) (%)	<i>g</i> (mm)	\bar{l}_1 (km) (%)	\bar{l}_2 (km) (%)	<i>TL</i> (km) (%)
S	0,40	2,55	0	–	–	0,40	2,57	–	–
				–	–		0,67	–	–
S	0,50	3,40	0	–	–	0,40	3,42	–	–
				–	–		0,44	–	–
S	0,40	1,40	0	–	–	0,40	1,53	–	–
S	0,50	1,50		–	–		9,07	–	–
S	0,32	0,20	0	–	–	0,32	0,318	–	–
S	0,40	0,90		–	–		59,0	–	–
S	0,50	1,50		–	–		–	–	–
S	0,63	0,50		–	–		–	–	–
S	0,40	1,45	0	–	–	0,40	1,57	–	–
S	0,50	0,75		–	–		8,28	–	–
S	0,63	0,50		–	–		–	–	–
S	0,90	0,50		–	–		–	–	–
S	0,40	1,30	0	–	–	0,40	1,40	–	–
S	0,50	1,25		–	–		7,69	–	–
S	0,63	0,50		–	–		–	–	–
S	0,32	0,20	0	–	–	0,32	0,318	–	–
S	0,40	0,60		–	–		59,0	–	–
S	0,90	4,0		–	–		–	–	–
S	0,40	0,75	1	0,806	–	0,40	0,806	–	1,34
BT	0,40	0,50		7,5	–		7,5	–	27,8
S	0,40	1,10		–	–		–	–	–
BT	0,40	0,50		–	–		–	–	–

6.2.1.3 Resultados LQ-SUITE

Para todos os resultados apresentados nesta seção, o desvio padrão σ apresenta valores pequenos, compatíveis com o da Tabela 6.2. Isso mostra que TIMEC apresenta intervalo de confiança que pode ser considerado aceitável para aplicação em sistemas DSL. Como foi descrito em [61], usar o valor do comprimento total pode melhorar significativamente a resposta obtida por TIMEC, mesmo que o valor da estimativa tenha erro de 10%. Mas para a obtenção do comprimento total, não havia nenhuma ferramenta para estimação desse parâmetro para a versão de TIMEC descrita no artigo. Já com o IRA esse parâmetro pode ser obtido facilmente e com grande precisão como mostrado na Tabela 6.3. Os resultados de TIMEC da Tabela 6.2 são mostrados na Tabela 6.5 de forma simplificada, onde são mostrados apenas os erros máximos de comprimento das seções para cada enlace, além do tempo de processamento. Como pode ser ver, comparando com Tabela 6.6, os resultados usando IRA são bem superiores com erros sempre menores que 5%. Além da estrutura topológica, bitolas e tipos das seções, ser sempre corretamente encontrada. Os tempos de processamento também são melhorados, apresentando valor máximo de 9 minutos, contra valores acima de 30 minutos para versão anterior de TIMEC.

Na Tabela 6.7 é mostrado os resultados obtidos pelo LQ-SUITE, ou seja, todos os módulos desenvolvidos trabalhando em conjunto. Como se pode ver, os resultados apresentaram erros zero, exceto para o enlace ETSI # 8. Para esse enlace a estrutura topológica é correta, incluindo as bitolas, com o erro máximo do comprimento das seções foi de 1,21 % e também apresentou valor de tempo de processamento maior que a versão com apenas TIMEC e IRA. Assim, para esse enlace, LQ-SUITE obteve resultados ligeiramente piores que o de TIMEC\IRA, da Tabela 6.6. Mas esse resultado ainda é bem superior ao resultado obtido apenas pela versão de TIMEC do artigo [61] e diante da qualidade das outras soluções, pode-se dizer que a versão do LQ-SUITE tem um desempenho de alta qualidade quando sob condições ideais, isto é, quando é usando dados simulados.

6.2.2 Dados Medidos

Na Tabela 6.8 são mostrados os enlaces usados nos testes usando dados medidos. Trata-se de enlaces de 1 a 3 seções e 2 enlaces com *bridged-taps*. Esses enlaces foram escolhidos buscando-se enlaces que poderiam ocorrer na prática e que estivesse dentro das possibilidades do laboratório de medição. Os procedimentos e o setup de medição são descritos no Apêndice A.

Tabela 6.5: Resultados obtidos aplicando somente TIMEC aos dados simulados dos enlaces ETSI. l_{max} (%) é o erro máximo de comprimento considerando todas as seções do enlace.

Enlace	Estrutura Topológica	l_{max} (%)	Tempo de processamento (min.)
ETSI#1	Correto	0,0	1
ETSI#2	Correto	0,0	1
ETSI#3	Correto	0,0	3
ETSI#4	Correto	0,005	20
ETSI#5	Correto	22,2	25
ETSI#6	Correto	1,59	8
ETSI#7	Correto	0,363	38
ETSI#8	Incorreto	45,5	36

Tabela 6.6: Resultados obtidos aplicando aplicando IRA e TIMEC aos dados simulados dos enlaces ETSI. l_{max} (%) é o erro máximo de comprimento considerando todas as seções do enlace.

Enlace	Estrutura Topológica	l_{max} (%)	Tempo de processamento (min.)
ETSI#1	Correto	0,016	< 1
ETSI#2	Correto	0,019	< 1
ETSI#3	Correto	0,0001	3
ETSI#4	Correto	0,39	6
ETSI#5	Correto	4,00	9
ETSI#6	Correto	0,00	3
ETSI#7	Correto	0,018	6
ETSI#8	Correto	0,126	8

6.2.2.1 Resultados do IRA

Na Tabela 6.9 é mostrado os resultados de IRA para dados medidos dos enlaces da Tabela 6.8. Como se pode ver, o IRA distinguiu corretamente enlaces seriais de enlaces com *bridged-tap*. Para esses últimos, os comprimentos das *bridged-taps* foram estimados com erro máximo de 6,2%. Para o enlace # 11, a bitola da *bridged-tap* foi incorretamente encontrada - 0,50 mm ao invés de 0,40 mm. Deve-se notar que esse valor de bitola poderá ser corrigido pelo TIMEC. Para o enlace #12, a bitola da *bridged-tap* foi corretamente encontrada. A posição da

Tabela 6.7: Resultados obtidos aplicando LQ-SUITE aos dados simulados dos enlaces ETSI. l_{max} (%) é o erro máximo de comprimento considerando todas as seções do enlace.

Enlace	Estrutura Topológica	l_{max} (%)	Tempo de processamento (min.)
ETSI#1	Correto	0,0	< 1
ETSI#2	Correto	0,0	< 1
ETSI#3	Correto	0,0	< 1
ETSI#4	Correto	0,0	3
ETSI#5	Correto	0,0	1
ETSI#6	Correto	0,0	< 1
ETSI#7	Correto	0,0	4
ETSI#8	Correto	1,21	12

bridged-tap não é considerada, já que esse valor é arbitrário conforme descrito na Seção 5.1.2. Considerando todos os enlaces, o erro máximo do comprimento de trabalho TL foi no máximo 5,9%, que pode ser considerado uma estimativa aceitável. Esse erro foi bem superior ao erro máximo obtido para dados simulados. Isso se deve as próprias condições de medição, sujeitas a interferências e ruídos, além de diferenças da velocidade de propagação teórica e a real, que podem afetar o valor estimado por IRA.

6.2.2.2 Resultados do FSF

O FSF também estimou corretamente o número de *bridged-taps*, fazendo a distinção correta entre os enlaces seriais e os de *bridged-taps*. O algoritmo encontrou corretamente a bitola da primeira seção para todos os enlaces. Para os enlaces com uma seção, o erro da primeira seção foi de apenas 6% e para todos esses enlaces o algoritmo indicou que eram constituídos de apenas uma seção, exceto para o enlace #3, que não pode ser confirmado como uma seção pelo FSF, conforme o critério definido na Seção 5.1.3. Para enlaces seriais de mais de uma seção, foram gerados dois valores para o comprimento da primeira seção. Um deles advém de um espúrio que não pôde ser eliminado pelo WFE. Assim ambos os valores são usados por TIMEC, conforme descrito na Seção 5.1.3. Para enlace #11, a posição e o comprimento da *bridged-tap* foram encontrados com erros aceitáveis, de 1,8% e 6% respectivamente. Pode-se ver que a *bridged-tap* vem logo após a primeira seção, já que o comprimento da primeira seção é igual a posição da *bridged-tap*. Já para o enlace #12, existem pelo menos duas seções antes da *bridged-tap*, já que o comprimento da primeira seção é menor que a posição da *bridged-tap*.

Tabela 6.8: Topologias dos enlaces teste para testar os métodos desenvolvidos usando dados medidos.

S indica um enlace serial e BT indica uma *bridged-tap*

Enlace #	Tipo	g (mm)	l (km)
1	S	0,40	0,40
2	S	0,40	0,50
3	S	0,40	1,0
4	S	0,50	0,40
5	S	0,50	0,50
6	S	0,50	1,0
7	S	0,40	0,50
	S	0,50	0,50
8	S	0,40	0,50
	S	0,50	1,0
9	S	0,40	1,0
	S	0,50	0,40
10	S	0,40	0,20
	S	0,50	1,0
	S	0,40	0,50
11	S	0,40	0,50
	BT	0,40	0,20
	S	0,50	1,0
12	S	0,40	0,20
	S	0,50	0,50
	BT	0,40	0,50
	S	0,50	0,50

O comprimento da *bridged-tap* não pode ser identificado, já que o pico positivo referente a *bridged-tap* não teve suas bordas identificadas pelo WFE.

Tabela 6.9: Resultado obtidos aplicando IRA aos dados medidos dos enlaces da Tabela 6.8. S indica um enlace serial, BT indica uma *bridged-tap* e TL trata-se do comprimento total.

Enlace #	ENLACES DE TESTE			ESTIMAÇÃO DO IRA							
	Tipo	g (mm)	l (km)	Tipo	g (mm)	\bar{l} (km)	E_l (%)	TL (km)	erro TL (%)		
1	S	0,40	0,40	Enlace Serial				0,416	4,0		
2	S	0,40	0,50	Enlace Serial				0,508	1,6		
3	S	0,40	1,0	Enlace Serial				1,06	6,0		
4	S	0,50	0,40	Enlace Serial				0,416	4,0		
5	S	0,50	0,50	Enlace Serial				0,508	1,6		
6	S	0,50	1,0	Enlace Serial				1,02	2,0		
7	S	0,40	0,50	Enlace Serial				1,02	2,0		
	S	0,50	0,50								
8	S	0,40	0,50	Enlace Serial				1,57	4,7		
	S	0,50	1,0								
9	S	0,40	1,0	Enlace Serial				1,48	5,71		
	S	0,50	0,40								
10	S	0,40	0,2	Enlace Serial				1,80	5,9		
	S	0,50	1,0								
	S	0,40	0,50								
11	S	0,40	0,50	S	0,40	0,785	–				
	BT	0,40	0,2	BT	0,50	0,208	4,0			1,57	4,67
	S	0,50	1,0	S	0,50	0,785	–				
12	S	0,40	0,2	S	0,40	0,624	–				
	S	0,50	0,50	–	–	–	1,25			4,17	
	BT	0,40	0,50	BT	0,40	0,531	6,2				
	S	0,50	0,50	S	0,50	0,624	–				

6.2.2.3 Resultados do LQ-SUITE

De novo aqui serão comparados os resultados de TIMEC com as informações obtidas por IRA em separado, onde essas informações *a priori* são o comprimento total e se o enlace

tem *bridged-taps*, com os resultados do LQ-SUITE, que envolve os três módulos, TIMEC, IRA e FSF. Como explicado anteriormente, esse teste é realizado para mostrar que apenas a informação do comprimento total e se o enlace tem *bridged-taps*, já produz respostas satisfatórias. No entanto, com o FSF as repostas apresentam ainda mais qualidade.

Na Tabela 6.11 são mostrados os resultados. Como pode-se ver, a resposta de TIMEC e IRA trabalhando em conjunto já obtêm corretamente a estrutura topológica (não considerando as bitolas) de todos os enlaces, com exceção do enlace #3, onde duas seções ao invés de uma foram encontradas. Para os enlaces #10 e #12, a bitola de uma das seções não foi corretamente encontrada. Apesar do erro do comprimento das seções chegar para alguns enlaces até 25,6%, o erro do comprimento total se manteve sempre abaixo de 5,5%. Isso se deve a compensações devido a variações no comprimento de outras seções. O erro do comprimento das *bridged-taps* foi de 6,5% e 5,2% para os enlaces #11 e #12, respectivamente. Para o LQ-SUITE, os erros dos comprimentos das seções foram sempre menores em comparação aos resultados anteriores até o enlace #9, com valor máximo 5,7%. Mas mesmo para os outros enlaces, houve equivalência de resultados, com ligeira vantagem para o LQ-SUITE. Para todos os enlaces, a estrutura topológica (sem incluir as bitolas) foi sempre corretamente encontrada, inclusive para o enlace #3. Apenas para os enlaces com *bridged-tap*, uma das bitolas foi incorretamente estimada, sendo nesses casos, a bitola da *bridged-tap*. O erro do comprimento total foi sempre abaixo de 2,4% considerando todos os enlaces, resultado melhor que o obtido anteriormente por TIMEC e IRA. O erro do comprimento das *bridged-taps* foi ligeiramente menor para o LQ-SUITE em comparação a versão sem o FSF, com valores de 5,5% e 5% para os enlaces #11 e #12, respectivamente.

Tabela 6.11: Resultados obtidos aplicando LQ-SUITE aos dados medidos dos enlaces da Tabela 6.8. O erro percentual é mostrado ao lado da estimativa. *S* indica um enlace serial, *BT* indica uma *bridged-tap* e *TL* trata-se do comprimento total.

Enlace #	TOPOLOGIA DE TESTE			TIMEC				LQ-SUITE			
	Tipo	g (mm)	l (km)	Tipo	g (mm)	\bar{l} (km)	E_l %	Tipo	g (mm)	\bar{l} (km)	E_l %
1	S	0,40	0,40	S	0,40	0,385	3,75	S	0,40	0,424	6,00
2	S	0,40	0,50	S	0,40	0,480	4,00	S	0,40	0,509	1,80
3	S	0,40	1,0	S	0,40	0,446	55,4	S	0,40	1,045	4,5
	-	-	-	S	0,50	0,519	-	-	-	-	-
4	S	0,50	0,40	S	0,50	0,386	3,50	S	0,50	0,403	0,75
5	S	0,50	0,50	S	0,50	0,477	4,60	S	0,50	0,508	1,80
6	S	0,50	1,0	S	0,50	0,963	3,70	S	0,50	1,018	1,80
7	S	0,40	0,50	S	0,40	0,403	19,4	S	0,40	0,486	2,80
	S	0,50	0,50	S	0,50	0,544	8,8	S	0,50	0,503	0,60
8	S	0,40	0,50	S	0,40	0,407	18,6	S	0,40	0,472	5,6
	S	0,50	1,0	S	0,50	1,012	1,2	S	0,50	0,992	0,8
9	S	0,40	1,0	S	0,40	0,925	7,5	S	0,40	0,943	5,7
	S	0,50	0,40	S	0,50	0,445	11,25	S	0,50	0,478	4,4
10	S	0,40	0,2	S	0,40	0,198	1,00	S	0,40	0,152	24,0
	S	0,50	1,0	S	0,50	1,039	3,9	S	0,50	0,990	1,0
	S	0,40	0,50	S	0,63	0,372	25,6	S	0,40	0,539	7,8
11	S	0,40	0,50	S	0,40	0,431	13,80	S	0,40	0,514	2,8
	BT	0,40	0,2	BT	0,40	0,187	6,5	BT	0,50	0,189	5,5
	S	0,50	1,0	S	0,50	1,006	0,60	S	0,50	0,966	3,4
12	S	0,40	0,2	S	0,40	0,191	4,50	S	0,40	0,194	3,0
	S	0,50	0,50	S	0,50	0,499	0,20	S	0,50	0,482	3,6
	BT	0,40	0,50	BT	0,50	0,474	5,20	BT	0,50	0,475	5,0
	S	0,50	0,50	S	0,50	0,462	7,60	S	0,50	0,500	0,0

6.3 Conclusões do Capítulo

Neste capítulo foram apresentados os resultados da metodologia de identificação de enlace proposta, mostrando que a metodologia é promissora. Os resultados obtidos foram melhores que os da técnica de identificação baseada em TDR recentemente descrita na literatura. Além disso, foi mostrado que as especializações da metodologia aumentaram a sua precisão além de reduzir significativamente o seu tempo de processamento, comprovando também que a metodologia é flexível e que pode facilmente estendida. Por fim, foram discutidos aspectos relacionados a precisão e desempenho dos modelos de linha.

Tabela 6.10: Resultados obtidos aplicando FSF aos dados medidos dos enlaces da Tabela 6.8. O erro percentual é mostrado ao abaixo da estimativa. *S* indica um enlace serial, *BT* indica uma *bridged-tap* e *TL* trata-se do comprimento total.

Enlace #	ENLACES DE TESTE			ESTIMAÇÃO DO FSF									
	Tipo	<i>g</i> (mm)	<i>l</i> (km)(%)	<i>nbt</i>	<i>posbt</i> (km)(%)	\bar{l}_{bt} (km)(%)	<i>g</i> (mm)	\bar{l}_1 (km)(%)	\bar{l}_2 (km)(%)	<i>TL</i> (km)(%)			
1	S	0,40	0,40	0	–	–	0,40	0,424	–	0,424			
								6	–	6			
2	S	0,40	0,50	0	–	–	0,40	0,509	–	0,509			
								1,8	–	1,8			
3	S	0,40	1,0	0	–	–	0,40	2,75	–	1,039			
								–	–	3,9			
4	S	0,50	0,40	0	–	–	0,50	0,403	–	0,403			
								0,75	–	0,75			
5	S	0,50	0,50	0	–	–	0,50	0,508	–	0,508			
								1,8	–	1,8			
6	S	0,50	1,0	0	–	–	0,40	1,018	–	1,018			
								1,8	–	1,8			
7	S	0,40	0,50	0	–	–	0,40	0,276	0,552	1,018			
	S	0,50	0,50					–	–	1,8			
8	S	0,40	0,50	0	–	–	0,40	0,276	0,445	1,527			
	S	0,50	1,0					–	–	1,8			
9	S	0,40	1,0	0	–	–	0,40	0,276	1,061	1,421			
	S	0,50	0,40					–	–	1,5			
10	S	0,40	0,2	0	–	–	0,40	0,233	0,424	1,718			
	S	0,50	1,0										
	S	0,40	0,50								–	–	1,05
11	S	0,40	0,50	1	0,509	0,212	0,40	0,509	–	–			
	BT	0,40	0,2										
	S	0,50	1,0								1,8	6,0	1,8
12	S	0,40	0,2	1	0,721	–	0,40	0,233	–	1,209			
	S	0,50	0,50										
	BT	0,40	0,50								3,0	16,5	0,75
	S	0,50	0,50										

Conclusão

O presente trabalho descreveu uma metodologia para identificação de topologias chamada LQ-SUITE que utiliza os novos recursos de medidas baseados nos padrões ITU-T G.992.3 e G.992.5 para ADSL 2 e ADSL 2+. A metodologia proposta usa ambas as medidas, SELT e DELT, e é baseada em otimização evolucionária multiobjetivo, otimização local e técnicas analíticas. O LQ-SUITE é constituído de três módulos, chamados de TIMEC, FSF e IRA. O TIMEC usa o parâmetro de espalhamento S_{11} e a função de transferência H , ambas medidas no domínio da frequência, para identificar a topologias de enlaces DSL através de um algoritmo genético (AG) multiobjetivo. Esse método utiliza modelos de linha e duas funções custos para avaliar as soluções geradas pelo AG, e assim aplicar o procedimento de seleção por ordenamento em frentes de Pareto, baseado na técnica NSGA-II. Essa técnica de otimização multiobjetivo foi utilizada por apresentar uma convergência melhor ou igual que outras técnicas baseada nas mesmas características de seleção por frentes de Pareto e elitismo, além de apresentar baixa complexidade computacional. O TIMEC foi comparado a uma técnica recente de identificação de topologias baseada em reflectometria no domínio do tempo, denominada neste trabalho de *SELT-tdr*. Os resultados apresentados mostraram que TIMEC apresentou melhores resultados comparados a essa técnica, com a identificação correta da estrutura topológica do enlace (número de seções e tipo de cada seção), além de apresentar erros pequenos na estimação dos comprimentos e bitolas para todos os enlaces testes utilizados nessa comparação. A técnica *SELT-tdr* depende de uma identificação muito precisa dos tempos de chegada de cada eco no reflectograma, de forma a estimar precisamente os comprimentos das seções associadas. Pequenos erros na estimação do comprimento das primeiras seções do enlace podem prejudicar de forma significativa o resultado do processo de identificação, apresentando topologias bem diferentes das topologias reais. Isto se explica pelo fato dessa técnica ser sequencial em que a estimação de uma seção depende da estimação da seção anterior. O que não ocorre com TIMEC, na qual são tratadas populações de soluções (enlaces) para obter a solução ótima do processo de identificação. Além disso, as técnicas de codificação\decodificação mostraram que TIMEC é flexível, aceitando diversos métodos baseados na quantidade e qualidade de informações pré-existentes sobre o enlace.

Dessa forma, quanto mais conhecimento se tiver da rede telefônica, maior a quantidade de inteligência que será possível de aplicar ao TIMEC, de forma rápida e simples. Um dos aspectos que podem prejudicar a utilização de algoritmos genéticos a problemas práticos, é o ajuste de seus parâmetros internos, como a probabilidade de cruzamento e mutação. Por isso, nessa tese foi desenvolvido um método que ajusta esses parâmetros automaticamente dependendo das características da evolução do algoritmo genético, melhorando a convergência do algoritmo. Assim, na prática o ajuste desses parâmetros será transparente aos técnicos responsáveis de operar a metodologia de identificação. Outra automatização aplicada, foi o desenvolvimento de uma ferramenta de medição que diminuí a interferência humana, realizando o controle dos equipamentos de medidas, a coleta dos dados, seu tratamento e gravação dos dados obtidos em disco. O FSF e o IRA que são as técnicas de pré-processamento utilizadas para obter informações *a priori* sobre o enlace, como o comprimento total, o número de *bridged-taps* do enlace e o comprimento de dessas *bridged-taps*, foram intensivamente testadas neste trabalho. O IRA realiza a análise da resposta impulsiva do enlace e FSF é aplicado a resposta TDR obtida de medidas de impedância de entrada. Com a aplicação dessas técnicas, a metodologia de identificação melhorou significativamente a sua precisão, além de ter seu tempo de processamento reduzido. Para dados simulados, os resultados foram exatos usando o LQ-SUITE, exceto para o enlace ETSI #8. Mas mesmo para esse enlace o erro máximo foi de apenas 1.21 %. Para dados medidos, os resultados foram aceitáveis e mostram que a metodologia, mesmo com o desvio do modelo de linha dos dados medidos, são suficientes precisos para serem usados pelos provedores de serviço DSL.

Como trabalho futuro, é possível o acréscimo de ferramentas de ajuste do modelo de linha, como a técnica descrita em [63]. Essa técnica seria aplicada em amostras de cabos da rede e seriam obtidos seus parâmetros elétricos e físicos, que minimizariam o desvio do modelo de linha e os dados medidos, aumentando a precisão do LQ-SUITE. Um exemplo pode ser dado no próprio laboratório de medição da UFPA, onde são usados cabos com uma cobertura de estanho em torno do núcleo de cobre. Assim, o valor da condutividade é diferente do valor da condutividade do cobre, que é usado no modelo VUB0. Como para altas frequências, a corrente tende a fluir mais na superfície devido ao efeito pelicular, a condutividade vista pela corrente tende a se afastar ainda mais da condutividade do cobre e se aproximar da condutividade do estanho, aumentando ainda mais o desvio do modelo de linha e os dados medidos. Com a técnica descrita em [63], que usa o modelo de altas frequência, o valor da condutividade que produz o melhor ajuste é o valor da condutividade do estanho. Dessa forma, usar esse valor melhoraria o ajuste do modelo e assim a precisão do LQ-SUITE. A partir de testes realizados e do que é descrito na literatura, O NSGA-II foi escolhido por apresentar melhores resultados que outras técnicas baseadas em elitismo por não dominância e também por ser

de fácil implementação. Mas como foi visto, apesar da técnica de otimização multiobjetivo escolhida ser de grande importância para a metodologia de identificação, existem outros aspectos que podem ser mais importantes, como as especializações realizadas que se baseiam nas condições específicas da identificação de enlaces DSL. Mesmo com essas considerações, outras técnicas de otimização multiobjetivo podem ser testadas, tanto as que já foram descritas na literatura quanto outras que surjam posteriormente. Como a metodologia foi desenvolvida de forma modular, a adaptação a outras técnicas pode ser realizada rapidamente, com a simples substituição de um módulo. Outro aspecto a ser explorado, é a utilização mais intensiva das estatísticas da rede onde a técnica está sendo aplicada. No trabalho atual algumas informações já são usadas, como a bitola 0,32 mm não ser aplicada em seções distantes do CO e ter apenas alguns metros de comprimento, já que essa bitola tem alta atenuação e por isso é muito improvável de ocorrer sob condições diferentes. Outra informação usada no algoritmo, é quanto ao número de *bridged-taps*. Baseado nos dados das redes brasileiras, descritas em [64], pode-se verificar que a existência de mais de 3 *bridged-taps* é muito improvável (abaixo de 2%). Assim, quando a informação do número de *bridged-taps* não pode ser detectada, os limites dessa variável, usada no processo de otimização, são de 0 a 3 *bridged-taps*. O método *typegauge* pode ser estendido para associar a escolha de uma bitola a uma determinada posição ao longo do enlace, baseado nos dados estatísticos de ocorrência daquela bitola nessa posição do enlace, similar ao que é feito em [65] e com dados estatísticos da rede em questão, como os descritos em [64] para as redes brasileiras. Isto pode ser realizado, fazendo-se um mapeamento não-uniforme dos M valores de bitolas disponíveis na faixa de $[0,1]$ a partir de M intervalos de comprimentos diferentes, que dependerão da probabilidade de ocorrência de uma determinada bitola na rede de telefonia em estudo. Assim, as bitolas mais prováveis de ocorrer em uma dada posição do enlace terão mais chances de serem escolhidas pelo método de identificação. Além do ajuste do modelo de linha e do uso da estatística da rede, é possível melhorar a interação das ferramentas de informação *a priori*, FSF e IRA. Se as informações de FSF forem aplicadas ao IRA, as estimativas desse último podem melhorar ainda mais, obtendo para o caso de enlaces com *bridged-taps*, quase todas as informações do enlace. O inverso também pode ser verdadeiro, com as informações de IRA sendo passadas ao FSF, melhorando também as suas estimativas. Dessa forma, para alguns casos, o TIMEC é aplicado apenas para ajuste fino da solução ou para confirmar a solução encontrada pelas ferramentas de informação *a priori*. Outra melhoria que poderia ser implementada, que apesar de secundária pode ser muito importante para viabilizar a utilização da ferramenta, é a implementação de uma interface para metodologia, onde todos os módulos poderiam ser controlados separadamente e os parâmetros importantes para o processo de estimativa podem ser mais facilmente visualizados e configurados.

Do que foi descrito, pode-se perceber que a metodologia possui precisão aceitável e

grande flexibilidade na identificação de enlaces DSL, que pode ser facilmente modificada e estendida para testar novas funcionalidades ou adequá-la às novas condições do processo de otimização.

Referências Bibliográficas

- [1] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II,” *IEEE Transactions on Evolutionary Computation*, vol. 6, pp. 182–197, Apr 2002.
- [2] T. Starr, J. M. Cioffi, and P. J. Silverman, *Understanding Digital Subscriber Line Technology*. Prentice-Hall, 1999.
- [3] J. J. Yoho, “Physically-Based Realizable Modeling and Network Synthesis of Subscriber Loops Utilized in DSL Technology,” Ph.D. dissertation, Virginia Polytechnic Institute and State University, 2001.
- [4] W. Y. Chen, *DSL: Simulation Techniques and Standards Development for Digital Subscriber Lines*. Alpel Publishing, 1997.
- [5] M. A. A. Medeiros, “Modelagem e Simulação Neuro-Fuzzy de um Controlador de Componente Reativa em Linha de Assinante ADSL,” Ph.D. dissertation, Centro Federal de Educação Tecnológica de Minas Gerais, 2006.
- [6] International Telecommunications Union - ITU, “ADSL2 Standard G.992.3,” 2003.
- [7] —, “ADSL2+ Standard G.992.5,” 2003.
- [8] D. Forum, “0adsl2 and adsl2plus - the new adsl standards,” 2003.
- [9] S. Galli and D. L. Waring, “Loop Make-up Identification via Single Ended Testing: Beyond Mere Loop Qualification,” *IEEE Journal on Selected Areas in Communications*, vol. 20, pp. 923–935, June 2002.
- [10] J. L. Fang, “Modeling and Characterization of Copper Access Systems,” Ph.D. dissertation, Stanford University, 2002.
- [11] S. Galli and K. J. Kerpez, “Single-Ended Loop Make-up Identification Part I: A method of analyzing TDR measurements,” *IEEE Transactions on Instrumentation and Measurements*, vol. 55, pp. 538–549, April 2006.

-
- [12] P. Boets, T. Bostoen, L. V. Biesen, and T. Pollet, "Preprocessing of Signals for Single-Ended Subscriber Line Testing," *IEEE Transactions on Instrumentation and Measurement*, vol. 55, pp. 1509 – 1518, 2006.
- [13] G. Long and J. Kamali, "Single-ended Line Probing Helps Speed up ADSL Mass Deployment," in *IIC-China/ESC China 2002*, 2002, pp. 57–60.
- [14] K. J. Kerpez and S. Galli, "Single-Ended Loop Make-up Identification Part II: Improved Algorithms and Performance Results," *IEEE Transactions on Instrumentation and Measurements*, vol. 55, pp. 538–549, April 2006.
- [15] T. Bostoen, P. Boets, M. Zekri, L. V. Biesen, T. Pollet, and D. Rabjins, "Estimation of the Transfer Function of a Subscriber by means of a One-Port Scattering Parameter Measurements at the Central Office," *IEEE Journal on Selected Areas in Communications*, vol. 20, pp. 936–948, June 2002.
- [16] T. Vermeiren, T. Bostoen, F. Louage, P. Boets, and X. O. Chebab, "Subscriber Loop Topology Classification by Means of Time Domain Reflectometry," *IEEE International Conference on Communications, Anchorage USA*, pp. 1998–2002, May 2003.
- [17] C. Neus, P. Boets, and L. V. Biesen, "Feature Estimation of One Port Scattering Parameters for Single Ended Line Testing," *XVIII Imeko World Congress*, 2006.
- [18] B. Celaya and D. E. Dodds, "Single-Ended DSL Line Tester," *Proceedings of IEEE Canadian Conference on Electrical and Computer Engineering, Niagara Falls, Canada*, pp. 2155–2158, May 2004.
- [19] D. Dodds, M. Shafique, and B. Celaya, "TDR and FDR Identification of Bad Splices in Telephone Cables," *IEEE Canadian Conference on Electrical and Computer Engineering, Ottawa, ON, Canada*, pp. 1031–1034, May 2006.
- [20] K. K. R. Faulkner and P. Golden, *Implementation and Applications of DSL Technology (Hardcover)*. USA: Auerbach Publications, 2007, chapter 5 – Evolving Test and Provisioning from POTS to xDSL Services.
- [21] J. L. Fang, C. Zeng, and J. Cioffi, "Bridged tap location estimation," *preprint*.
- [22] M. Nilsson, "Management of a DSL Copper Network using Built-in Loop Qualification Tools," Ph.D. dissertation, Royal Institute of Technology, 2005.
- [23] R. E. Matrick, *Transmission Lines For Digital and Communication Networks-An Introduction to Transmission Lines, High-Frequency and High-speed Pulse Characteristics and Applications*. IEEE, 1993.

-
- [24] D. Rauschmayer, *ADSL/VDSL Principles: A Practical and Precise Study of Asymmetric Digital Subscriber Lines and Very High Speed Digital Subscriber Lines*. Macmillan Technology Series, 1998.
- [25] ANSI, “(t1e1.4/98-043r3 - Very-high-speed Digital Subscriber Lines, System Requirements,” 1998.
- [26] R. van den Brink, *Cable Reference Models for Simulating Metallic Access Networks*. ETSI STC. TM6, 1998.
- [27] D. Franklin, J. Chicharo, and J. Xi, “A new channel model for adsl and vdsl systems,” in *School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, NSW, Australia, Fifth International Symposium on Signal Processing and its Applications, ISSPA '99*, Aug. 1999.
- [28] J. W. Cook, R. H. Kirkby, M. G. Booth, K. T. Foster, D. E. A. Clarke, and G. Young, “The noise and crosstalk environment for adsl and vdsl systems,” *IEEE Comm. Mag.*, pp. 73–78, 1999.
- [29] P. Boets, “Frequency domain identification of transmission lines from time domain measurements,” Ph.D. dissertation, Vruce Universiteit Brussel, 1997.
- [30] P. Boets and L. V. Biesen, “Metallic 2-Wire Parametric Line Models - a Survey,” in *XVIII Imeko World Congress - Metrology for a Sustainable Development*, 2006.
- [31] ANSI, “T1.413 - Network to Customer Installation Interface - Asymmetric Digital Subscriber Lines (ADSL) Metallic Interface,” 2004.
- [32] ETSI-TS 101 270-1 V1.2.1, “Transmission and Multiplexing (TM); Access Transmission Systems on Metallic Access Cables; Very high-speed Digital Subscriber Line (VDSL). Part 1: Functional Requirements,” 1999.
- [33] L. Heylen and J. Musson, *Cable Models Predict Physically Impossible Behaviour in Time Domain*. ETSI STC. TM6, 1999.
- [34] T. Nordström and S. Trautmann, *Proposal for Parametric ADSL Cable Models*. ETSI STC. TM6, 2002.
- [35] T. Nordström, *A model for an Austrian PE04 Cable*. ETSI STC. TM6, 2003.
- [36] V. Belevitch, “Theory of the Proximity Effect in Multiwire Cables - Part I,” *Philips Research Reports*, vol. 32, pp. 16–43, 1977.

-
- [37] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA: Addison-Wesley, 1989.
- [38] F. Herrera, M. Lozano, and J. L. Verdegay, “Tackling real-coded genetic algorithms: Operators and tools for behavioural analysis,” *Artificial Intelligence Review*, vol. 12, 1998.
- [39] Z. Michalewicz, *Genetic Algorithms Plus Data Structures Equals Evolution Programs*, 1994.
- [40] E. A. Amorim, “Fluxo de potência Ótimo em sistemas multimercados através de um algoritmo evolutivo multiobjetivo,” Ph.D. dissertation, Faculdade de Engenharia de Ilha Solteira - UNESP, 2006.
- [41] R. I. Haupt and S. E. Haupt, *Practical Genetic Algorithms*, 2nd ed. Wiley, 1998.
- [42] K. Deb, *Multi-Objective Optimization using Evolutionary Algorithms*. John Wiley & Sons, 2001.
- [43] T. Blickle, “Theory of evolutionary algorithms and application to system synthesis,” Ph.D. dissertation, Swiss Federal Institute of Technology, 1996.
- [44] L. Davis, *Handbook of Genetic Algorithms*. VNR Comp. Library, 1990.
- [45] A. S. Catarina, “Aplicações de algoritmos genéticos em sistemas de informações geográficas,” Ph.D. dissertation, INPE, 2004.
- [46] S. Ventura, C. Romero, A. Zafra, J. Delgado, and C. Hervás, “Jelec: a java framework for evolutionary computation,” *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, vol. 12, no. 4, pp. 381–392, February 2008. [Online]. Available: <http://dx.doi.org/10.1007/s00500-007-0172-0>
- [47] T. Bäck, F. Hoffmeister, and H. P. Schwefel, “A survey of evolution strategies,” in *In Lashon B. Belew, Richard K.; Booker, editor, Proceedings of the 4th International Conference on Genetic Algorithms*, 1991, pp. 2–9.
- [48] H. M. Voigt and T. Anheyer, “Modal mutations in evolutionary algorithms,” *Proc. IEEE Int. Conf. on Evolutionary Computation*, vol. I, pp. 88–92, 1994.
- [49] W. G. C. Ticona, “Aplicação de algoritmos genéticos multi-objetivo para alinhamento de sequências biológicas,” Ph.D. dissertation, Instituto de Ciências Matemáticas e de Computação - ICMC-USP, 2003.

-
- [50] C. A. C. Coello, "An empirical study of evolutionary techniques for multiobjective optimization in engineering design," Ph.D. dissertation, Department of Computer Science, Tulane University, 2006.
- [51] J. D. Schaffer, "Multiple Objective Optimization with Vector Evaluated Genetic Algorithms," in *Genetic Algorithms and their Applications: Proceedings of the First International Conference on Genetic Algorithms*, 1985, pp. 93–100.
- [52] C. M. Fonseca and P. J. Fleming, "Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization," in *Proceedings of the Fifth International Conference on Genetic Algorithms*, 1993, pp. 416–423.
- [53] E. Zitzler, M. Laumanns, and L. Thiele, "SPEA2: Improving the Strength Pareto Evolutionary Algorithm," Computer Engineering and Networks Laboratory (ETH Zurich), Tech. Rep. 103, Dec. 2001.
- [54] J. Horn, N. Nafpliotis, and D. E. Goldberg, "A niched pareto genetic algorithm for multiobjective optimization," in *Proceedings of the First IEEE Conference on Evolutionary Computation, IEEE World Congress on Computational Intelligence*, vol. 1, 1994, pp. 82–87.
- [55] E. Zitzler, K. Deb, and L. Thiele, "Comparison of Multiobjective Evolutionary Algorithms: empirical results (Revised Version)," Computer Engineering and Networks Laboratory (ETH Zurich), Tech. Rep. 70, Dec. 1999.
- [56] N. Srinivas and K. Deb, "Multi-objective function optimization using non-dominated sorting genetic algorithm," *Evolutionary Computation*, vol. 2, pp. 221–248, 1994.
- [57] K. D. Tran, "Elitist Non-dominated Sorting GA-II (NSGA-II) as a Parameter-less Multi-objective Genetic Algorithm," in *Proc. IEEE SoutheastCon*, 2005.
- [58] International Telecommunications Union - ITU, "Test Procedures for DSL Transceivers Standard - G.996.1," 1999.
- [59] V. D. Lima, A. B. R. K. Jr, G. P. S. Cavalcante, J. C. W. A. Costa, and K. Ericsson, "Feature extraction on loop make up identification using wavelet-based singularity detection in time domain," *A ser submetido*.
- [60] L. Team, "UFA05 LQ Seventh Quarterly Report (Q7)," Federal University of Pará (UFPA), Tech. Rep., Aug. 2008.

-
- [61] C. Sales, R. M. Rodrigues, F. Lindqvist, J. Costa, A. Klautau, K. Ericson, J. R. i Riu, and P. O. Börjesson, “Line Topology Identification using Multi-Objective Evolutionary Computation,” *Accepted for publication in IEEE Transactions on Instrumentation & Measurement*.
- [62] Ph.D. dissertation.
- [63] G. S. Borges, R. R. Menezes, C. S. Sales, J. Costa, and K. Ericson, “Cable Parameters Identification for DSL Systems,” *A ser submetido*.
- [64] Telebrás, “Asymmetric Digital Subscriber Lines (ADSL): aspectos de transmissão na rede metálica de assinantes,” 1998.
- [65] K. J. Kerpez and S. Galli, “Improved Algorithms for Single-Ended Loop Make-Up Identification,” *IEEE International Conference on Communications, 2006. ICC '06.*, vol. 1, pp. 67–72, June 2006.

Publicações Relacionadas

- Conferências:
 - C. S. Sales Jr., L. V. de Souza, R. M. Rodrigues, A. L. S. Castro, G. P. S. Cavalcante, J. C. W. Costa and J. R. i Riu. “VDSL Systems: Measurements and Analyzes,” SPIE Optics East 2007, Boston-MA, USA.
 - C. S. Sales Jr., L. D. S. Alcantara, R. M. Rodrigues, A. L. S. Castro, J. C. W. Costa, F. Lindqvist and J. R. i Riu. “Single-Ended Transfer Function Estimation of Telephone Links for Deployment of xDSL Services,” XXII Simpósio brasileiro de telecomunicações - SBrT’05, 04-08 de setembro de 2005, Campinas, SP.

- Revista:
 - C. S. Sales, R. Menezes, K. Ericson, J. R. I. Riu, F. Lindqvist, A. Klautau and J. C. W. Costa, “Loop Topology Identification using Multi-Objective Evolutionary Computation via Two-Port Measurements,” *Instrumentation and Measurement, IEEE Transactions* on Volume 58, Issue 11, Page(s):1 - 1

- Patente:
 - C. S. Sales, J. R. I. Riu, F. Lindqvist and J. C. W. Costa, “Method and arrangement for loop qualification in a digital subscriber line (DSL) system.”2008.

Apêndice A

Setup de Medição de Função de Transferência *H* e Parâmetro de Espalhamento S_{11}

A seguir será descrita a campanha de medição para obtenção das medidas usadas na Seção 6.2.2. É mostrados os equipamentos e os cabos reais utilizados, além de uma breve descrição do processo de medição.

A.1 Lista de Equipamentos Utilizados

- Dois baluns *North Hills* 0301BB, com as seguintes especificações:
 - Frequência de corte inferior: 10 kHz;
 - Frequência de corte superior: 60 MHz;
 - Impedância de suas portas: lado UNB $50\ \Omega$, lado BAL $100\ \Omega$.
- Padrões de calibração: aberto, curto, *load* e *through*;
- Agilent 4395A (analisador de rede);
- Agilent 87512A (conjunto de teste de transmissão/reflexão);
- Cabos com conectores BNC de 1,10 metros de comprimento;
- Cabos de par trançado, cujas características estão listadas na Tabela A.12 e mostrados na Figura A.1.

Tabela A.12: Main features of the employed cables.

Nome de referência	Diâmetro (mm)	Comprimento (m)	Número de pares	Fabricante
LQ 31	0.4	200	16	Ericsson
LQ 38	0.4	500	16	Ericsson

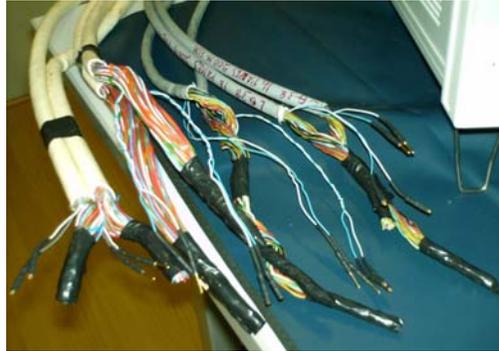


Figura A.1: Cables used during the measurement process.

A.2 Configuração dos Equipamentos de Medição

- Frequência inicial: 4.3125 kHz;
- Frequência final: 2,208 MHz;
- Número de pontos: 512.
- Largura de banda de resolução: 300 Hz.

A.3 Breve Descrição do Processo de Medição

Dois enlaces foram selecionados para essa campanha de medição. As quantidades medidas para cada enlace foram o parâmetro de espalhamento S_{11} com a porta remota em aberto e, função de transferência H . Cada medida foi realizada na faixa de frequência do ADSL2+ e repetidas cinco vezes de forma a calcular as estatísticas dos dados medidos de cada enlace.

O processo de calibração realizado para o parâmetro de espalhamento S_{11} foi o 1-port (open/short/100 Ω - carga resistiva com 1% de precisão). Para a função de transferência foi o *response calibration (open/short/through)*.

O *setup* de medição utilizando para a função de transferência H e parâmetro de espalhamento S_{11} é mostrado na Figura A.2, destacando-se que para o caso do parâmetro de espalhamento S_{11} apenas um balun foi utilizado.



Figura A.2: *Setup* de medição para função de transferência H e parâmetro de espalhamento S_{11} , sendo que para esse último apenas um balun é utilizado deixando a porta de saída em aberto.

As medidas de parâmetros de espalhamento e função de transferência H estão disponíveis em:

- <https://www2.laps.ufpa.br/protected/measurements/S11/LQSELT> – DELT_No.02.2007, and;
- <https://www2.laps.ufpa.br/protected/measurements/TrF/LQSELT> – DELT_No.01.2007.

Apêndice B

Enlaces ETSI

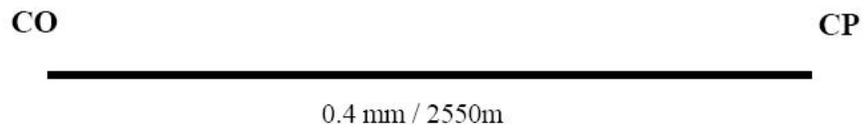


Figura B.1: ETSI#1.

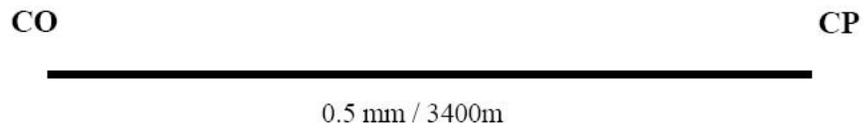


Figura B.2: ETSI#2.

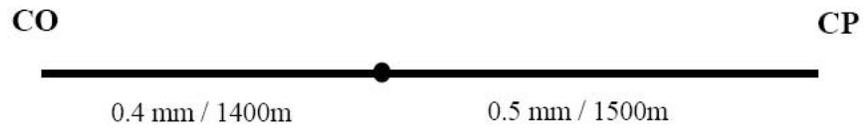


Figura B.3: ETSI#3.

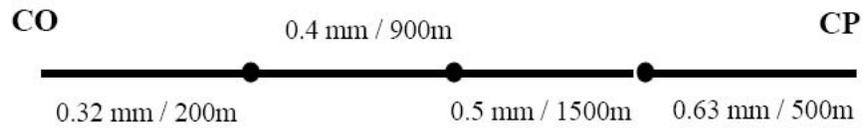


Figura B.4: ETSI#4.

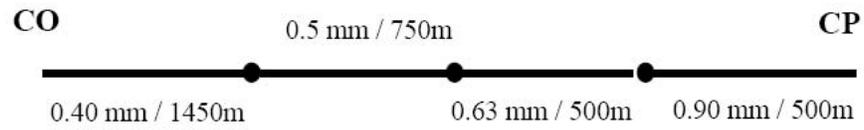


Figura B.5: ETSI#5.

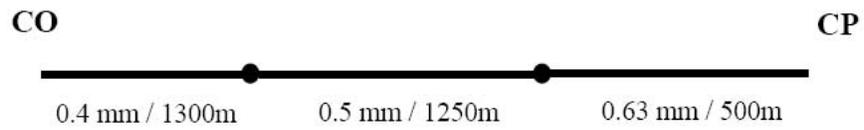


Figura B.6: ETSI#6.

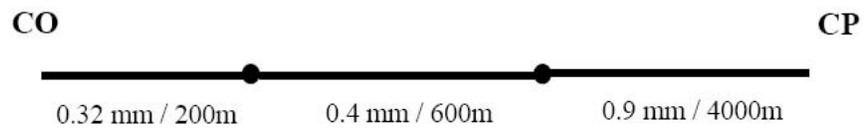


Figura B.7: ETSI#7.

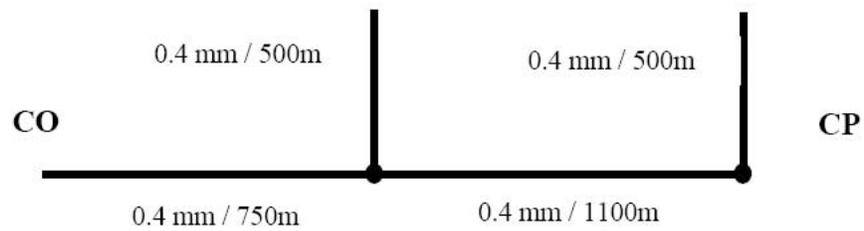


Figura B.8: ETSI#8.