



UNIVERSIDADE FEDERAL DO CEARÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELEINFORMÁTICA

Wendley Souza da Silva

NEUROCOMPUTAÇÃO NA PREDIÇÃO DE CONTORNOS DE TRÁFEGO
DE BACKBONE INTERNET

FORTALEZA - CEARÁ
FEVEREIRO - 2010

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Wendley Souza da Silva

NEUROCOMPUTAÇÃO NA PREDIÇÃO DE CONTORNOS DE TRÁFEGO DE
BACKBONE INTERNET

Dissertação

Dissertação submetida ao corpo docente da
Coordenação do Programa de Pós-Graduação em
Engenharia de Teleinformática da Universidade
Federal do Ceará como parte dos requisitos
necessários para obtenção do grau de MESTRE EM
ENGENHARIA DE TELEINFORMÁTICA.

Área de concentração: Sinais e sistemas

Danielo Gonçalves Gomes

(orientador)

FORTALEZA - CEARÁ
FEVEREIRO - 2010

S584n Silva, Wendley Souza da
Neurocomputação na predição de contornos de tráfego de Backbone
internet / Wendley Souza da Silva, 2010.
106 f. ;il. enc.

Orientador: Prof. Dr. Danielo Gonçalves Gomes

Área de concentração: Sinais e Sistemas

Dissertação (mestrado) - Universidade Federal do Ceará, Centro de
Tecnologia. Departamento de Engenharia de Teleinformática, Fortaleza,
2010.

1. Redes neurais (computação) 2. Telecomunicações-tráfego 3.
Algoritmos I. Gomes, Danielo Gonçalves (orient.) II. Universidade Federal
do Ceará – Pós-Graduação em Engenharia de Teleinformática III.Título

CDD 621.38

WENDLEY SOUZA DA SILVA

NEUROCOMPUTAÇÃO NA PREDIÇÃO DE CONTORNOS DE TRÁFEGO DE
BACKBONE INTERNET

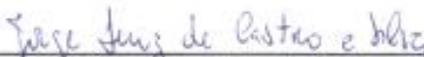
Dissertação submetida à Coordenação do Programa de Pós-Graduação em Engenharia de Teleinformática, da Universidade Federal do Ceará, como requisito parcial para a obtenção do grau de Mestre em Engenharia de Teleinformática. Área de concentração SINAIS E SISTEMAS.

Aprovada em 25/02/2010.

BANCA EXAMINADORA



Prof. Dr. Danielo Gonçalves Gomes (Orientador)
Universidade Federal do Ceará -UFC



Prof. Dr. Jorge Luiz de Castro e Silva
Universidade Estadual do Ceará -UECE



Prof. Dr. Guilherme de Alencar Barreto
Universidade Federal do Ceará -UFC



Prof. Dr. Miguel Franklin de Castro
Universidade Federal do Ceará -UFC

*Dedico este trabalho aos meus pais
Francisco de Assis e Fátima Souza,
minha irmã Francisca Karen,
minha noiva Aline de Carvalho
e ao Matemático Eliowald
Carvalho (in memoriam),
pelo constante apoio e incentivo*

Agradecimentos

Ao meu orientador, Prof. Danielo Gonçalves Gomes, a quem sou grato pela orientação, paciência e confiança depositada.

Ao professor Gabriel Antoine Louis Paillard, coordenador do Curso de Eng^a da Computação, UFC-Sobral, pela confiança e apoio ofertados.

Aos professores e funcionários do Departamento de Engenharia de Teleinformática que, direta ou indiretamente, participaram do desenvolvimento deste trabalho.

Ao professor Guilherme de Alencar Barreto, pelo apoio durante esta jornada.

"A sabedoria começa na reflexão." (Sócrates)

Sumário

Lista de Tabelas	x
Lista de Figuras.....	xi
Lista de Símbolos.....	xiii
Lista de Siglas	xiv
Resumo	xv
<i>Abstract</i>	xvi
Introdução.....	1
1.1 Contexto e Motivação.....	1
1.2 O problema	2
1.3 Proposta resumida	3
1.4 Contribuições	3
1.5 Organização do documento	3
Revisão Bibliográfica	5
2.1 Modelagem e caracterização de tráfego de rede	5
2.1.1 Auto-similaridade e LRD.....	7
2.2 Técnicas clássicas de predição	9
2.2.1 Séries temporais.....	9
2.2.2 Médias móveis – MA.....	10
2.2.3 Alisamento exponencial	10
2.3 RFC 2988.....	13
2.4 Conclusão	13
Proposta – algoritmos neurais para predição de contornos	14
3.1 Extração de contornos por máximo de grupos.....	14
3.1.1 Janela fixa	15
3.1.2 Janela variável	17
3.1.3 Critérios de escolha do método	18
3.2 Redes neurais aplicadas aos máximos de grupos	20
3.2.1 Justificativa	20
3.2.2 Rede Perceptron Multicamadas.....	21

3.2.6	Topologia NARX.....	27
3.3	Conclusão	33
	Resultados e discussão.....	35
4.1	Caracterização e metodologia da simulação	35
4.2	Métricas analisadas.....	37
4.3	Resultados.....	37
4.3.1	Cenário 1: MA com dados originais.....	38
4.3.2	Cenário 2: MA com contornos por ECJF	39
4.3.3	Cenário 3: MA com contornos por ECJV.....	40
4.3.4	Cenário 4: AES com dados originais	41
4.3.5	Cenário 5: AES com contornos por ECJF.....	43
4.3.6	Cenário 6: AES com contornos por ECJV	44
4.3.7	Cenário 7: HW com dados originais	45
4.3.8	Cenário 8: HW com contornos por ECJF.....	46
4.3.9	Cenário 9: HW com contornos por ECJV	48
4.3.10	Cenário 10: RFC 2988 com dados originais	49
4.3.11	Cenário 11: RFC 2988 com contornos por ECJF	50
4.3.12	Cenário 12: RFC 2988 com contornos por ECJV	51
4.3.13	Cenário 13: MLP com dados originais	53
4.3.14	Cenário 14: MLP com contornos por ECJF	54
4.3.15	Cenário 15: MLP com contornos por ECJV	55
4.3.16	Cenário 16: Narx com dados originais.....	56
4.3.17	Cenário 17: Narx com contornos por ECJF	58
4.3.18	Cenário 18: Narx com contornos por ECJV.....	59
4.3.19	Conjunto de simulações com os dados originais	61
4.3.20	Conjunto de simulações com os contornos (ECJF).....	63
4.3.21	Conjunto de simulações com os contornos (ECJV)	65
4.4	Análise dos resultados	67
4.5	Conclusão	69
	Considerações Finais.....	71
5.1	Limitações do trabalho.....	72
5.2	Trabalhos futuros	72

Referências bibliográficas	73
Apêndices	78
Apêndice A – Código-fonte do algoritmo ECJF	78
Apêndice B – Código-fonte do algoritmo ECJV	79
Apêndice C – Código-fonte do algoritmo MLP	81
Apêndice D – Código-fonte do algoritmo NARX	86

Lista de Tabelas

Tabela 1 - Disposição dos agrupamentos a1 e a2 com uso de janela fixa sem superposição	16
Tabela 2 - Disposição dos agrupamentos a1 , a2 e a3 com uso janela variável e erros admissível máximo de 12 unidades	17
Tabela 3 - Determinação da JAG (<i>traces</i> CAIDA, 2000 pontos).....	19
Tabela 4 - Determinação da qtde. de neurônios ocultos na rede MLP (ETA = 0,01 e m = 0)... 23	23
Tabela 5 - MSE e MAC da simulação neural com MLP, alternando-se a taxa de aprendizagem.....	25
Tabela 6 - MSE da simulação neural com MLP, alternando-se o valor do termo de momento	26
Tabela 7 - Determinação da qtde. de neurônios ocultos na rede NARX (ETA = 0,01 e m = 0). 29	29
Tabela 8 - MSE e MAC da simulação neural com NARX, alternando-se a constante de aprendizagem.....	31
Tabela 9 - MSE da simulação neural com NARX, alternando-se o valor de momentum	32
Tabela 10 - Resumo da configuração da rede MLP	34
Tabela 11 - Resumo da configuração da rede NARX.....	34
Tabela 12 - Cenários de teste.....	36
Tabela 13 - MSE das simulações utilizando-se dados originais.....	61
Tabela 14 - MAC das simulações utilizando-se dados originais.....	61
Tabela 15 - MSE das simulações utilizando-se contornos por ECJF	63
Tabela 16 - MAC das simulações utilizando-se contornos por ECJF	63
Tabela 17 - MSE das simulações utilizando-se contornos por ECJV	65
Tabela 18 - MAC das simulações utilizando-se contornos por ECJV.....	65
Tabela 19 - MSE médio dos métodos preditores.....	67
Tabela 20 - MAC médio dos métodos preditores	68
Tabela 21 - Tempo (em segundos) para execução de 10 rodadas dos algoritmos neurais	69

Lista de Figuras

Figura 1 - Exemplo de contorno obtido pelo algoritmo ECJF.....	15
Figura 2 - Arquitetura de uma MLP com uma camada escondida (adaptado de Aragão Jr., 2008) .	21
Figura 3 - Curvas das Figuras de mérito usadas na determinação do número de neurônios na rede MLP	24
Figura 4 - Curvas das Figuras de mérito usadas na determinação da taxa de aprendizagem na rede MLP	26
Figura 5 - Curvas das Figuras de mérito usadas na determinação do termo de momento na rede MLP	27
Figura 6 - Arquitetura de uma NARX com D entradas	28
Figura 7 - Curvas das Figuras de mérito usadas na determinação do número de neurônios na camada escondida da rede NARX	30
Figura 8 - Curvas das Figuras de mérito usadas na determinação da constante de aprendizagem na rede NARX.....	32
Figura 9 - Curvas das Figuras de mérito usadas na determinação do termo de momento na rede NARX	33
Figura 10 - Metodologia aplicada.....	36
Figura 11 - Cenário 1 - Método MA com os dados originais normalizados	38
Figura 12 - Cenário 1 - Histograma do erro.....	38
Figura 13 - Cenário 2 - Método MA com os contornos obtidos por ECJF	39
Figura 14 - Cenário 2 - Histograma do erro.....	40
Figura 15 - Cenário 3 - Método MA com os contornos obtidos por ECJV.....	40
Figura 16 - Cenário 3 - Histograma do erro.....	41
Figura 17 - Cenário 4 - Método AES com os dados originais	42
Figura 18 - Cenário 4 - Histograma do erro.....	42
Figura 19 - Cenário 5 - Método AES com contornos obtidos por ECJF	43
Figura 20 - Cenário 5 - Histograma do erro.....	43
Figura 21 - Cenário 6 - Método AES com contornos obtidos por ECJV.....	44
Figura 22 - Cenário 6 - Histograma do erro.....	45
Figura 23 - Cenário 7 - Método HW com os dados originais	45
Figura 24 - Cenário 7 - Histograma do erro.....	46
Figura 25 - Cenário 8 - Método HW com contornos obtidos por ECJF	47
Figura 26 - Cenário 8 - Histograma do erro.....	47
Figura 27 - Cenário 9 - Método HW com contornos obtidos por ECJV.....	48
Figura 28 - Cenário 9 - Histograma do erro.....	48
Figura 29 - Cenário 10 - Método RFC 2988 com os dados originais	49
Figura 30 - Cenário 10 - Histograma do erro.....	50
Figura 31 - Cenário 11 - Método RFC 2988 com contornos obtidos por ECJF	50
Figura 32 - Cenário 11 - Histograma do erro.....	51
Figura 33 - Cenário 12 - Método RFC 2988 com contornos obtidos por ECJV.....	52

Figura 34 - Cenário 12 - Histograma do erro.....	52
Figura 35 - Cenário 13 - Método MLP com os dados originais	53
Figura 36 - Cenário 13 - Histograma do erro.....	53
Figura 37 - Cenário 14 - Método MLP com contornos obtidos por ECJF	54
Figura 38 - Cenário 14 - Histograma do erro.....	55
Figura 39 - Cenário 15 - Método MLP com contornos obtidos por ECJV.....	55
Figura 40 - Cenário 15 - Histograma do erro.....	56
Figura 41 - Cenário 16 - Método Narx com os dados originais	57
Figura 42 - Cenário 16 - Histograma do erro.....	57
Figura 43 - Cenário 17 - Método Narx com contornos obtidos por ECJF.....	58
Figura 44 - Cenário 17 - Histograma do erro.....	58
Figura 45 - Cenário 18 - Método Narx com contornos obtidos por ECJV	59
Figura 46 - Cenário 18 - Histograma do erro.....	60
Figura 47 - Média do MSE dos valores preditos para dados originais	62
Figura 48 - Média dos MAC dos valores preditos para dados originais.....	62
Figura 49 - Média do MSE dos valores preditos para contorno por ECJF	64
Figura 50 - Média do MAC dos valores preditos para contorno por ECJF	64
Figura 51 - Média do MSE dos valores preditos para contorno por ECJV.....	66
Figura 52 - Média do MAC dos valores preditos para contorno por ECJV.....	66
Figura 53 - MSE médio dos métodos preditores.....	67
Figura 54 - MAC médio dos métodos preditores	68

Lista de Símbolos

\mathbb{R}	conjunto dos números reais
x	variável escalar, i.e., $x \in \mathbb{R}$
t	índice indicativo de tempo contínuo
ρ	autocorrelação
E	o valor esperado
k	defasagem
μ	média
σ	desvio padrão.
σ^2	variância amostral
X	série de dados
Σ	somatório
n	índice indicativo de tempo discreto
\mathbb{Z}	conjunto dos números inteiros
\hat{x}	variável escalar predita
S	variável de sazonalidade
b_1	sinal base ou componente permanente
b_2	componente linear
S_t	fator de sazonalidade aditiva
ϵ_t	componente de erro aleatório
y	dados observados
p	período da componente sazonalidade
θ	vetor com os parâmetros de alisamento
k	constante para obtenção do RTT predito
a	agrupamentos
w_{ij}	peso associado à ligação entre entrada j e neurônio i da camada intermediária de uma rede supervisionada
m	termo de momento

Lista de Siglas

GBPS	Gigabit por segundo
MSE	<i>Mean Square Error</i>
MAC	Métrica das Amplitudes Conjugadas
QoS	<i>Quality of Service</i>
IEEE	<i>Institute of Electrical and Electronics Engineers</i>
IETF	<i>Internet Engineering Task Force</i>
P2P	<i>Peer-to-peer</i>
ISP	<i>Internet Service Providers</i>
MPLS	<i>Multiprotocol Label Switching</i>
ISP	<i>Internet service provider</i>
RFC	<i>Request for Comments</i>
TCP	<i>Transmission Control Protocol</i>
VDSP	<i>Video on Demand Service Providers</i>
VoIP	<i>Voice over Internet Protocol</i>
LRD	<i>Long Range Dependences</i>
ACF	<i>Autocorrelation Function</i>
LAN	<i>Local Area Network</i>
WAN	<i>Wide Area Network</i>
SRD	<i>Short Range Dependence</i>
MA	<i>Moving Average</i>
AES	Alisamento Exponencial Simples
HW	Holt-Winters
RTT	<i>Round Trip Time</i>
ECJF	Algoritmo Extrator de Contornos com Janelas Fixas
ECJV	Algoritmo Extrator de Contornos com Janelas Variáveis
MLP	<i>MultiLayer Perceptron</i>
NARX	<i>Nonlinear Autoregressive with eXogenous input</i>
JAG	Janela de Agrupamento
CAJ	Constante de ajuste
ETA	Constante de aprendizagem
CRP	Critério com relação ponderada
SLA	<i>Service Level Agreement</i>

Resumo

Diante do avanço tecnológico das redes de telecomunicações, com modernas aplicações multimídia, torna-se imprescindível o uso de processos de modelagem analítica, de simulação numérica e de medições práticas para concepção, planejamento e análise quantitativa de seus desempenhos. Neste contexto, classificação e predição de tráfego de redes constituem um objeto relevante de estudo. Apesar de diversos métodos e modelos de predição de tráfego de rede disponíveis na literatura, nota-se que a maioria efetua uma estimativa direta a partir de um grande volume de dados de entrada. Nesta dissertação são avaliados dezoito cenários de predição de contornos de traços de tráfego reais de *backbone* Internet de 10 Gbps. Contornos são as representações máximas do tamanho de grupos de pacotes de tráfego, determinados através de um algoritmo extrator, proposto neste trabalho. Propomos uma nova abordagem para a predição, de tráfego de rede, na medida em que são utilizados modelos conexionistas (neurais) cujas entradas são os contornos e não as amostras pontuais dos traços de tráfego. Os experimentos realizados foram analisados à luz de duas métricas: a clássica MSE (*Mean Square Error*) e uma métrica aqui proposta, chamada MAC (Métrica das Amplitudes Conjugadas). Os resultados obtidos indicam que realizar neuropredição de tráfego de redes a partir de seus contornos é bastante promissor.

Palavras-chave: algoritmo ECJF, algoritmo ECJV, auto-similaridade, contornos, métrica MAC, MLP, modelagem de tráfego, NARX, predição.

Abstract

Regarding the progress in telecommunications networks development, mainly due to the mobile and modern multimedia applications, the use of analytical modeling process, simulation and measurement, are essential to a better performance analysis. The observation of network traffic is at the core of many fundamental network operation and maintenance activities, such as the enforcement of quality of service guarantees or traffic engineering. Several works on network traffic prediction propose a direct estimation of the network traces, requiring a large amount of input data. In this work, we analyze eighteen scenarios of a real Internet backbone link (OC192) traffic traces. We propose a new approach for the network traffic prediction, where neural predictors use the traffic trace max-group values. The experiments were realized based on two metrics: the classical MSE (Mean Square Error) and another metric here proposed, named MAC (Métrica das Amplitudes Conjugadas, in Portuguese). The results show that our approach is able to predict Internet backbone traffic trace files with high accuracy.

Keywords: ECJF algorithm, ECJV algorithm, self-similarity, outline, MAC measurement, MLP, traffic models, NARX, prediction.

Capítulo 1

Introdução

1.1 Contexto e Motivação

A evolução das redes de telecomunicações, bem como as exigências das suas modernas aplicações e aplicativos multimídia, sintetizando o conceito de qualidade de serviço (*Quality of Service, QoS*), têm tornado indispensável a utilização de ferramentas de modelagem analítica, de simulação numérica e de medições práticas para concepção, planejamento e análise quantitativa de seus desempenhos.

Em particular, o crescimento da Internet na última década tem sido fortemente caracterizado por mudanças consideráveis nos modos pelos quais os usuários finais comportam-se, interagem entre si e utilizam a rede. Pesquisadores da área, órgãos de padronização (e.g. *Institute of Electrical and Electronics Engineers, IEEE* e *Internet Engineering Task Force, IETF*) e operadoras de rede em geral respondem a essas mudanças projetando, padronizando e implantando complexos arcabouços e arquiteturas para medições e classificação de tráfego. De fato, questões relativas a tráfego de rede estão no cerne das atividades de pesquisa e desenvolvimento para garantias de QoS, engenharia de tráfego, detecção de anomalias e prevenção de ataques em rede.

Entretanto, o advento das novas categorias de aplicações multimídia interativas tais como jogos em rede e *peer-to-peer* (p2p), o aumento da presença de tráfego malicioso e o amplo uso de algoritmos de criptografia impõem obstáculos consideráveis às tarefas de medição, análise e classificação do tráfego na Internet.

Um grande desafio aos pesquisadores e engenheiros da área é o desenvolvimento de novos algoritmos capazes de superar as limitações das técnicas tradicionais de inspeção dos pacotes, baseadas somente na análise das portas nos campos de seus cabeçalhos. Outros obstáculos majoritários ao progresso das pesquisas nessa área são a premissa da garantia da privacidade dos usuários e a necessidade de redução da quantidade de tráfego coletada, armazenada e processada.

Neste contexto, classificação e predição de tráfego constituem um objeto relevante enquanto tema de pesquisa atual bem como uma prática estratégica para médios e grandes provedores de acesso (*Internet Service Providers*, ISP). Estudos de novas técnicas podem ser aplicados, por exemplo, em controle de congestionamento (Li, Zhao, Xu, & Liu, 2005), controle de admissão de conexões (Shu, Zhigang, Jidong, & Yang, 2000), no domínio de engenharia de tráfego, implementação de túneis MPLS (*Multiprotocol Label Switching*) entre os pares de roteadores de uma rede (Cortez, Rio, Rocha, & Sousa, 2006).

1.2 O problema

As propriedades inerentes do tráfego das redes em geral, e da Internet em particular, apresentam-se como obstáculos na predição de seus próximos “passos” com vistas a uma alocação de recursos sob demanda para minimização de desperdícios.

O crescimento permanente do uso da Internet e o aumento da capilaridade dos acessos aos usuários finais geram a necessidade de melhores práticas na gerência dos recursos de rede com escalabilidade através de novos estudos e soluções de engenharia para modelagem de tráfego.

A maioria dos métodos de predição existentes efetua uma análise global dos dados. Tais métodos podem levar à perda de detalhes (*underfitting*) ou à inclusão de ruído (*overfitting*), por tratar os dados de forma indistinta. Os traços de tráfego apresentam um grande volume de informações, dificultando ou protelando uma predição a partir dos dados originais (traços de tráfego).

O problema tratado nesta dissertação é a predição de uma série representativa do tráfego real Internet.

1.3 Proposta resumida

A proposta consiste na concepção e implantação de um preditor de contornos de tráfego Internet baseado em técnicas conexionistas. Contornos são os valores máximos de um grupo de pacotes de tráfego – tais contornos são determinados através de um algoritmo extrator com uso de janela fixa ou variável, aqui também proposto. Em particular, os traços (arquivos do tipo *log*) utilizados nos experimentos desta dissertação são dados de tráfego real de um *backbone*¹ Internet, a ser descrito em maiores detalhes no Capítulo 4.

1.4 Contribuições

Considerando as lacunas dispostas na Seção 1.2, concernentes à modelagem e predição de tráfegos de redes IP, as principais contribuições deste trabalho envolvem:

- Análise do desempenho das técnicas clássicas de predição estatística frente a traços de tráfego reais e atualizados de um *backbone* Internet;
- Proposição de um algoritmo extrator de contornos de tráfego a fim de reduzir o volume dos dados de entrada e prover uma abordagem de predição em diferentes níveis de escala;
- Proposição de um modelo conexionista para predição de tráfego de rede a partir de traços originais e de contornos;
- Proposição de métrica de comparação baseada na amplitude dos tráfegos.

1.5 Organização do documento

¹Espinha dorsal de um provedor Internet (ISP), onde as vazões de tráfego são elevadas – atualmente atingem 40 Gbps.

O restante desta dissertação está organizado em quatro capítulos. Um breve comentário sobre cada um deles é visto a seguir.

No Capítulo 2 é feita uma revisão bibliográfica dos modelos de tráfego de redes bem como das técnicas clássicas de predição estatística investigadas nesta dissertação.

No Capítulo 3 é detalhada a proposta desta dissertação, demonstrando a abordagem conexionista para contornos de traços de redes.

No Capítulo 4 são apresentados os resultados obtidos a partir dos métodos preditivos discutidos nos Capítulos 2 e 3 com as técnicas de extração de contornos explicitadas no Capítulo 3, perfazendo 18 cenários possíveis.

No Capítulo 5 são discutidas as principais conclusões desta dissertação e sugestões para trabalhos e pesquisas futuros que podem dar continuidade à mesma.

Ao final do manuscrito, um apêndice traz os códigos desenvolvidos para os algoritmos propostos no Capítulo 3.

Capítulo 2

Revisão Bibliográfica

Este é um capítulo de referencial teórico tratando de modelagem e caracterização de tráfego de rede (Seção 2.1), de técnicas clássicas de predição estatística (Seção 2.2) investigadas nesta dissertação e do padrão Internet RFC (*Request for Comments*) 2988 (Seção 2.3).

2.1 Modelagem e caracterização de tráfego de rede

Desde meados da década de 1990, vários trabalhos sobre modelagem e medições de tráfego de rede têm afirmado que a complexidade inerente das suas informações não são satisfatoriamente caracterizadas por modelos tradicionais Poissonianos (Leland, Taqqu, Willinger, & Wilson, 1994), (Paxson & Floyd, 1995) e (Orosz, Krauskopf, & Wilson, 2005). Modelos alternativos, tais como os que levam em consideração a dependência de longa duração do tráfego (Bianchi, 2005), bem como ferramentas e técnicas de avaliação e controle de tráfego acuradas têm sido objeto de pesquisa nesta área (Xie, Wang, & Wang, 2007).

A predição de tráfego de redes é de significativo interesse em diversos domínios (Brakmo & Peterson, 1995), (Roychoudhuri & Al-Shaer, 2005), (Gomes, Agoulmine, Bennani, & de Souza, 2007), (Soysal & Schmidt, 2010) e (Aragão Jr. & Barreto, 2010). Brakmo & Peterson (1995) propõem o algoritmo TCP Vegas visando proporcionar ao TCP (*Transmission Control Protocol*) uma maior taxa de transferência de dados e diminuir a perda de segmentos na rede em períodos de congestionamento. Em seu trabalho Roychoudhuri & Al-Shaer (2005) propõem um *framework* para prever a perda de pacotes e congestionamento com base em variações de atraso fim-a-fim, tornando possível a recuperação proativa de erro e evitando congestionamentos. Gomes, Agoulmine,

Bennani, & de Souza (2007) estabelecem melhorias em um módulo de predição neural em aplicações de vídeo sob demanda (VDSP - *Video on Demand Service Providers*). Soysal & Schmidt (2010) empregam aprendizado de máquina para classificação de diferentes tipos de tráfego de Internet, incluindo os oriundos de aplicativos de jogos e p2p. Aragão Jr. & Barreto (2010) estabelecem uma abordagem adaptativa para a predição de atrasos em aplicações VoIP (*Voice over Internet Protocol*) usando vários modelos lineares e neurais, tais como MLP e Elman.

De acordo com Zhani, Elbiaze, & Kamoun (2010) distinguem-se duas categorias de predição: predições de longo e de curto período. Predições de longo período provêm uma detalhada previsão da carga de trabalho e padrões de tráfego permitindo avaliar futuros requisitos de capacidade. Predições de curto período é pertinente para alocação dinâmica de recursos, podendo ser usado na melhoria da qualidade de serviço (QoS), bem como mecanismos de congestionamento e controle de recursos através da adaptação dos parâmetros de rede para as características do tráfego, e encaminhamento de pacotes.

Classicamente, a modelagem de tráfego em telefonia convencional e em redes de computadores se baseou, ao longo de vários anos, quase que exclusivamente nos modelos poissonianos ou, mais genericamente, markovianos (Carvalho, 1998). Contudo, estudos mais recentes mostram uma significativa mudança na compreensão dos tráfegos de redes, apresentando propriedades como auto-similaridade e dependência temporal longa – LRD (*Long Range Dependences*) (Wang, Li, & Zhao, 2007).

Previsão de tráfego tem sido amplamente investigada desde a aceitação da auto-similaridade e da característica LRD de tráfego de redes (Leland, Taqqu, Willinger, & Wilson, 1994), (Paxson & Floyd, 1995), (Abry & Veitch, 1998), (Karagiannis, Molle, & Faloutsos, 2004) e (Gomes, Agoulmine, Bennani, & de Souza, 2007). Contudo, apesar dessa característica poder causar intensos efeitos sobre o desempenho da rede em termos de perda e atraso (Zhani, Elbiaze, & Kamoun, 2008), vários estudos têm demonstrado que a auto-similaridade pode ser explorada para caracterizar ou para prever o tráfego, a fim de controlar os recursos de rede (Papagiannaki, Taft, Zhang, & Diot, 2003), (Owezarski & Larrieu, 2004) e (Scherrer, Larrieu, Borgnat, Owezarski, & Abry, 2006).

Cortez *et al.* (2007) indicam que atualmente está-se voltando a atenção para predição de tráfegos de backbones, ao invés de um enlace (*link*) específico. Em seus trabalhos, utilizaram-se de uma abordagem neural na predição de tráfego TCP/IP de um backbone, contudo, suas propostas foram superadas por outros métodos.

O tráfego típico de dados contém rajadas seguidas por longos períodos de silêncio (amortização) na proporção de 1:1000 ou até mesmo de 1:10000 em comparação com

uma razão de 1:3 para tráfego de voz (Kleinrock, 2002). Essa natureza particular do tráfego de dados motivou o desenvolvimento das técnicas baseadas no princípio do compartilhamento dinâmico dos recursos da rede (roteadores, concentradores, enlaces, entre outros), nos quais são multiplexados em um único canal, não de forma determinística com reserva de tempo, por exemplo, mas de forma aleatória, de modo que o acesso à rede seja imediato a qualquer instante e com qualquer duração. Em seu trabalho Gomes, Agoulmine, Bennani, & de Souza (2007) alertam que os multiplexadores sujeitos a processos com LRD podem apresentar perdas massivas de pacotes.

2.1.1 Auto-similaridade e LRD

Série temporal é uma coleção de observações $(x_1, x_2, x_3, \dots, x_t)$ sequenciais ao longo do tempo t . Vale notar também que o tempo pode ser substituído por outra variável como espaço ou profundidade.

Pode-se caracterizar a dependência temporal (i.e. memória) entre observações em diferentes tempos pela avaliação da função de autocorrelação (ACF – *Autocorrelation Function*) que é $\rho(k)$. ACF mede a similaridade entre uma série X_t e a sua versão deslocada X_{t+k} :

$$\rho(k) = \frac{E[(X_t - \bar{x})(X_{t+k} - \bar{x})]}{\sigma^2}, \quad (1)$$

$$\bar{x} = E[x_t], \quad (2)$$

$$\sigma^2 = E[(X_t - \bar{x})^2], \quad (3)$$

em que E , k , \bar{x} e σ representam, respectivamente, o valor esperado, defasagem, média e desvio padrão.

Em processos de tráfego de redes de computadores, geralmente X refere-se ao número de pacotes e *bytes* em intervalos de tempo consecutivos, significando que X descreve o volume de *bytes*/pacotes observados em um enlace no intervalo de tempo t .

A dependência de longa duração é uma importante característica do tráfego, possuindo relevantes implicações em ferramentas teóricas para análise e projetos de sistemas

computacionais, tal como teoria de filas (Vieira & Ling, 2008). É essencial que os modelos de tráfego capturem as características de natureza auto-similar para que se possa gerar indicadores de desempenho de redes ou fazer previsões confiáveis de seu desempenho.

O estudo da caracterização de tráfego em redes tem sido abordado na literatura segundo duas análises principais: a dos modelos Poissonianos e dos modelos fractais ou auto-similares. Leland, Taqqu, Willinger, & Wilson (1994) mostraram a característica auto-similar do tráfego em redes Ethernet baseado em rigorosas análises estatísticas efetuadas nos laboratórios da Bellcore (Leland & Wilson, 1991) com graus diferentes de auto-similaridade dependente da carga na rede. Especificamente, observaram que a variabilidade do tráfego Internet não era variante nas escalas de tempo observadas, ou seja, o tráfego não era suavizado com a agregação, conforme indicado pelo modelo poissoniano. A partir de então e durante toda a década de 90, uma série de trabalhos estudaram e confirmaram a presença de auto-similaridade tanto em tráfegos de rede locais (LAN – *Local Area Network*) como nos de rede de longa distância (WAN – *Wide Area Network*), tais como Paxson & Floyd (1995) e Crovella & Bestavros (1996).

Um processo tem LRD quando sua função de autocorrelação apresenta um decaimento lento, à medida que k aumenta $\sum_{k=1}^{\infty} |\rho(k)| = \infty$; em contraste aos modelos de tráfego poissonianos. Juntos, a dependência de longo prazo e a invariância a escalas de tempo são padrões típicos de um processo estocástico auto-similar, conceito matemático introduzido originalmente por Mandelbrot (Mandelbrot & Ness, 1968). Em suma, a auto-similaridade é uma propriedade referente à invariância da distribuição para qualquer incremento do processo (Beran, 1994), ou quando a série temporal apresenta a mesma estrutura de autocorrelação para diferentes níveis de agregação (Leland, Taqqu, Willinger, & Wilson, 1994).

Seja $\{x(t), t \in T\}$ um processo estocástico. $\{x(t), t \in T\}$ é estritamente estacionário se e somente se $(x(t_1), x(t_2), \dots, x(t_n))$ e $(x(t_1 + k), x(t_2 + k), \dots, x(t_n + k))$ possuem a mesma distribuição conjunta $\forall k, n \in \mathbb{Z}$.

Um processo estocástico é fracamente estacionário (ou estacionário de segunda ordem) se e somente se $\forall t \in T, E(x(t)) = u, V(x(t)) = \sigma^2$ e $\gamma(t, t + k) = \gamma(k)$. A esperança e a variância independem do tempo e a covariância depende apenas da diferença de tempos.

Um processo estritamente estacionário é chamado de auto-similar com parâmetro de auto-similaridade $0 < H < 1$ se $m^{1-H} \times x^{(m)} \stackrel{\text{def}}{=} x, m = 1, 2, \dots$, isto é, a distribuição de probabilidade é a mesma, tanto para a soma agregada ($x^{(m)}$) quanto para a série original (x). Se x é auto-similar, então x e $x^{(m)}$ tem a mesma função de autocorrelação.

A auto-similaridade de segunda ordem captura a propriedade de que a estrutura de correlação é exatamente ou assintoticamente preservada pelo processo agregado.

Processos auto-similares de segunda ordem são caracterizados por um decaimento hiperbólico a ACF e utilizados amplamente em modelos LRD (Karagiannis, Molle, & Faloutsos, 2004). Inversamente, um rápido decaimento caracteriza dependência de curto alcance (*Short Range Dependence*, SRD). Dessas definições, infere-se que uma série é caracterizada como LRD se $0,5 < H < 1$. Valores de H próximos ou tendendo a 1, indicam forte dependência temporal.

2.2 Técnicas clássicas de predição

Predição ou previsão é um problema relevante para várias ciências, senão todas. Os métodos variam de acordo com o problema e dados disponíveis. Devido a dificuldade em modelar problemas reais, surgiram técnicas que visam simplificar e auxiliar a previsão em séries temporais que descrevem tais problemas. As principais técnicas de previsão são a modelagem estatística, a modelagem de sistemas inteligentes ou ainda uma modelagem híbrida utilizando ambas as técnicas.

Dentre as técnicas clássicas de predição de séries temporais as tratadas nesta dissertação estão as de Médias Móveis (MA), o modelo de Alisamento Exponencial Simples (AES) e a técnica de Holt-Winters.

2.2.1 Séries temporais

Uma série temporal é de tempo contínuo quando as observações quando as observações podem ser medidas em qualquer instante de tempo. Definindo o conjunto $T = \{t : t_1 < t < t_2\}$ a série temporal será denotada por $\{x(t) : t \in \mathbb{R}\}$. Uma série temporal é dita ser de tempo discreto quando as observações são feitas em tempos específicos, geralmente equiespaçados. Definindo o conjunto $T = \{t_1, \dots, t_n\}$ a série temporal será denotada por $\{x_t : t \in T\}$. Por simplicidade, pode-se fazer $T = \{1, 2, 3, \dots, n\}$.

2.2.2 Médias móveis – MA

Os modelos de tráfegos auto-regressivos definem a próxima observação na sequência como uma função explícita de observações anteriores dentro de uma janela definida entre valores atuais e passados. Como exemplo deste modelo está o método de previsão por médias móveis MA (*Moving Average*) que utiliza como previsão para um determinado ponto futuro as médias das observações passadas. Este método fornece uma previsão de curto prazo (*short-terms*), o que o torna uma interessante escolha quando os dados não apresentam tendência ou sazonalidade; contudo, existem meios para se aplicar MA a séries que possuam tendência e/ou sazonalidade (Makridakis, Wheelwright, & Hyndman, 1998).

As médias móveis se apresentam como simples, centradas ou ponderadas. Os modelos de médias simples, adotados neste trabalho, podem ser expressas como

$$y_t = \hat{x}_t = \frac{x_{t-1} + x_{t-2} + \dots + x_{t-n}}{n}, \quad (4)$$

em que n representa o tamanho da janela de tempo considerada para determinar a média e x_t o valor predito.

O termo Média Móvel é utilizado porque à medida que a próxima observação se torna disponível, a média das observações é recalculada, incluindo essa observação no conjunto de observações e desprezando a observação mais antiga (Morettin & Tolo, 1981).

2.2.3 Alisamento exponencial

2.2.3.1 Alisamento Exponencial Simples – AES

O método conhecido como Alisamento Exponencial Simples se assemelha ao da Média Móvel por diminuir das observações da série temporal o comportamento aleatório pelo alisamento dos dados históricos. Entretanto, a inovação introduzida pelo Alisamento Exponencial Simples advém do fato de este método atribuir pesos diferentes a cada observação da série. Enquanto que na Média Móvel as observações usadas para encontrar a previsão do valor futuro contribuem em igual proporção para o cálculo dessa previsão, no Alisamento Exponencial Simples as informações mais recentes são evidenciadas pela aplicação de um fator que determina essa importância (Wheelwright & Makridakis, 1985).

São especialmente úteis como um meio para suavizar (filtrar) oscilações aleatórias que existem com as seguintes propriedades: (i) redução do peso é atualizado em um dado antigo; (ii) é extremamente fácil computar; e (iii) quantidade mínima de dados é requerida (Makridakis, Wheelwright, & Hyndman, 1998). O argumento para o tratamento diferenciado das observações da série temporal é fundamentado na suposição de que as últimas observações contêm mais informações sobre o futuro e, portanto, são mais relevantes para a previsão. Em Gardner Jr. (2006) são descritas formulações, propriedade e apresentados os principais estudos sobre o método AES. Gardner Jr. & Diaz-Saiz (2008) apresentam o uso do método AES em dados provenientes de telecomunicações.

De acordo com Morettin & Tolo (2004) pode ser descrito pela seguinte equação:

$$\hat{Z}_t = aZ_t + (1 - a)\hat{Z}_{t-1}, \quad Z_0 = Z_1, \quad t = 1, \dots, N \quad (5)$$

sendo a a constante de alisamento, e Z_t valor exponencialmente alisado. Se a equação for resolvida recursivamente, pode-se observar que valores antigos de Z_t exercem menor influência sobre a previsão \hat{Z}_t .

2.2.3.2 Método de Holt-Winters

Uma variável nova a considerar quando se trata de métodos de previsão para séries temporais, é a sazonalidade, que se denota por S . S indica de quantas em quantas observações a sazonalidade pode ser observada, e estimada a partir de experimentos, sendo crucial para o método de previsão. Este modelo possui dois tipos de procedimentos, multiplicativo e aditivo, cuja utilização depende das características das séries consideradas.

Morettin & Tolo (2004) comentam que as vantagens desse modelo são o fácil entendimento e aplicação não dispendiosa, adequada para séries com padrão de comportamento mais geral; e que as desvantagens são dificuldades de determinar os valores mais apropriados das constantes de suavização e/ou impossibilidade de determinar as propriedades estatísticas, tais como média e variância da previsão e, conseqüentemente, a construção de um intervalo de confiança. Existem dois modelos: Holt-Winters Aditivo e Multiplicativo. O primeiro é utilizado quando a amplitude da variação sazonal mantém-se constante, ou seja, a diferença entre o maior e o menor ponto de demanda nos ciclos permanece constante com o passar do tempo. Holt-Winters Multiplicativo é utilizado quando a amplitude da variação sazonal aumenta com o tempo, ou seja, a diferença entre o maior e o menor ponto de demanda nos ciclos cresce com o passar do tempo.

Uma grande desvantagem deste método clássico é que ele pode ser fortemente influenciado pela presença de *outliers* na série temporal (Croux, Gelper, & Fried, 2008). Isso pode ser superado, incorporando idéias da literatura em técnicas de alisamento, como discutido em Cipra (1992). Kotsialos, Papageorgiou, & Poulimenos (2005) apresentam um uso prático da metodologia Holt-Winters em dados de vendas de duas empresas alemãs.

Sua forma geral é dada por:

$$y_t = b_1 + b_2 t + S_t + \epsilon_t \quad (6)$$

em que b_1 é o sinal base, também chamado de componente permanente, b_2 é o componente linear, S_t é o fator de sazonalidade aditiva, e ϵ_t é o componente de erro aleatório.

Este método produz resultados similares ao Alisamento Exponencial Linear, sendo, no entanto, capaz de manipular séries temporais que além de apresentarem tendência nos dados, apresentam também sazonalidade (Wheelwright & Makridakis, 1985).

O método Holt-Winters Aditivo, adotado nesta dissertação, é usualmente definido pelas seguintes equações (Bermúdez, Segura, & Vercher, 2005):

$$a_i = \alpha(y_i - c_{i-p}) + (1 - \alpha)(a_{i-1} + b_{i-1}) \quad (7)$$

$$b_i = \beta(a_i - a_{i-1}) + (1 - \beta)b_{i-1} \quad (8)$$

$$c_i = \gamma(y_i - a_{i-1} - b_{i-1}) + (1 - \gamma)c_{i-p} \quad (9)$$

em que $\{y_i\}_{i=1}^n$ são os dados observados, p é o período da componente sazonalidade, e $\theta = (\alpha, \beta, \gamma)$ é o vetor com os parâmetros de alisamento.

2.3 RFC 2988²

O RFC 2988 data do ano 2000, e foi originalmente pensado para estimar o tempo de viagem de ida e volta (RTT – *Round Trip Time*) no protocolo de congestionamento TCP (Paxson & Allman, 2000). Diante da simplicidade deste método, realizamos testes de viabilidade e o adotamos neste trabalho, com fim na predição de série temporal.

Com base na análise de Allman & Paxson (1999), Paxson & Allman (2000) concluíram que o tempo de retransmissão TCP não deve ser inferior a um segundo.

A Equação abaixo ilustra o método de estimação proposto no RFC 2988:

$$RTT_{\text{estimado}} = (1 - k) * RTT_{\text{estimado}} + k * RTT_{\text{amostra}} \quad (10)$$

a qual está escrita sob a forma de um comando de linguagem de programação. Com valor de $k = 0,125$ (Kurose & Ross, 2006), da Equação 11 obtém-se a seguinte equação:

$$RTT_{\text{estimado}} = 0,875 * RTT_{\text{estimado}} + 0,125 * RTT_{\text{amostra}} \quad (11)$$

2.4 Conclusão

Uma importante característica no tráfego com dependência temporal é que as observações vizinhas são muito correlacionadas, residindo, então, o interesse em analisar e modelar esta dependência.

Foram apresentadas algumas características de processos auto-similares e discutidas suas relações com LRD e o tráfego de redes.

Dentre as técnicas clássicas de predição descritas neste capítulo, encontram-se o método das Médias Móveis (MA), o modelo de Alisamento Exponencial Simples (AES) e o modelo de Holt-Winters.

O padrão Internet RFC 2088 embora originalmente desenvolvido para a estimativa de RTT, foi adotado nesta dissertação na predição de séries temporais.

² Disponível em <http://www.ietf.org/rfc/rfc2988.txt>

Capítulo 3

Proposta – algoritmos neurais para predição de contornos

Almejando a redução do número de informações a serem lançadas à entrada dos mecanismos preditores neurais e estatísticos (doravante denominada simplesmente de redução), as séries originais dos traços de tráfego deveriam passar por um processo de compactação sem perder a significância das informações. Diante disso, propõe-se a extração de contornos desses traços, através de dois algoritmos de extração: o algoritmo Extrator de Contornos com Janelas Fixas (ECJF) e com Janelas Variáveis (ECJV).

Propõe-se ainda a predição de traços originais e seus contornos com o uso dos modelos conexionistas MLP (*MultiLayer Perceptron*) e NARX (*Nonlinear Autoregressive with exogenous input*). Todos os códigos dos algoritmos aqui propostos estão disponíveis no apêndice ao final deste manuscrito.

3.1 Extração de contornos por máximo de grupos

Considera-se, nesta dissertação, que contorno de um grupo (intervalo) de valores é definido pelo seu valor máximo. Desta forma, a formação de um contorno se dá por extrações sucessivas do máximo de grupos de valores.

Na Figura 1 tem-se a ilustração gráfica de um contorno obtido pelo algoritmo ECJF utilizando-se do traço Bellcore³.

³ Traços obtidos do laboratório *Bellcore Morristown Research and Engineering*, clássicos na literatura, disponíveis em <http://ita.ee.lbl.gov/html/contrib/BC.html>

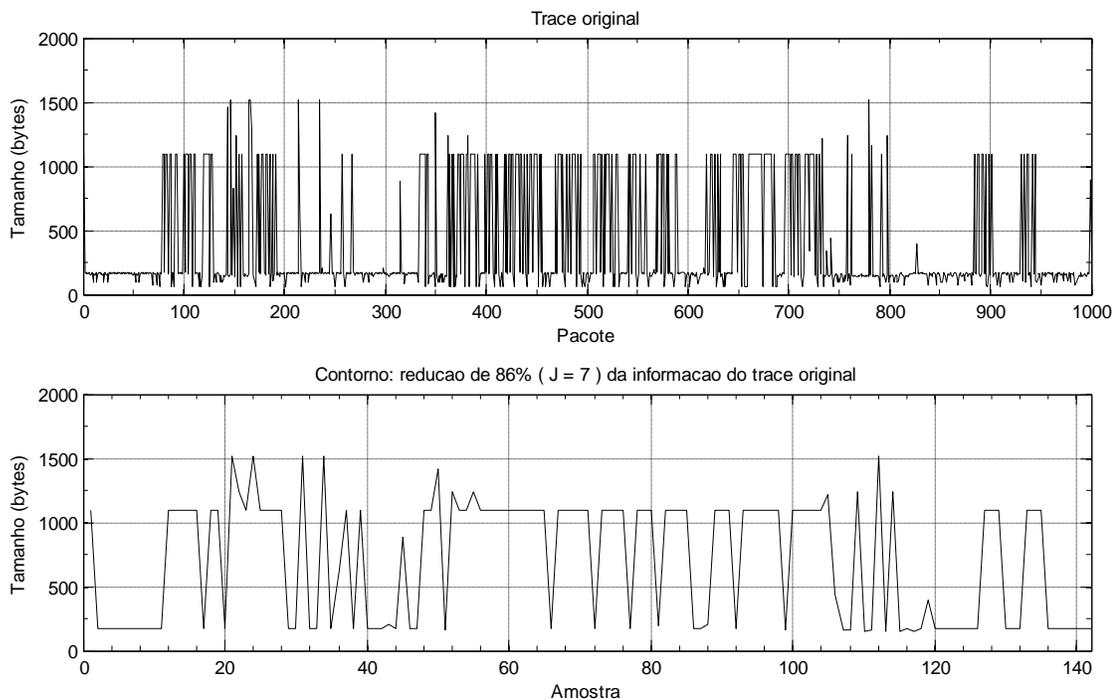


Figura 1 - Exemplo de contorno obtido pelo algoritmo ECJF

Nas duas subseções seguintes demonstra-se a forma de atuação do algoritmo relacionando-se com o fator “janela”, que pode ser fixa ou variável em tempo de execução.

3.1.1 Janela fixa

O fator determinante para se obter elevada taxa de redução é o tamanho da Janela de Agrupamento (JAG). Chamamos de janela o intervalo de atuação do algoritmo em uma determinada etapa de processamento. Considerando uma janela de tamanho igual a 5 e a seguinte sequência de observações

$$\{2; 7; 2; 5; 3; 8; 4; 4; 3; 2\}$$

tem-se que será possível a abrangência dos dados com o uso da janela a_1 na posição 1 em um tempo t_1 , e em seguida a janela a_2 na posição 6 em um tempo t_2 , através do deslocamento de 5 unidades, que é o tamanho da respectiva janela, conforme observado na Tabela 1.

Tabela 1 - Disposição dos agrupamentos $a1$ e $a2$ com uso de janela fixa sem superposição

Janela	$a1$					$a2$				
Posição	1	2	3	4	5	6	7	8	9	10
Valor	2	7	2	5	3	8	4	4	3	2

Dessa forma, o algoritmo Extrator de Contornos com uso de Janela Fixa sem superposição (ECJF) necessita do tamanho da janela como um parâmetro, de maneira que retorne uma boa representatividade dos contornos da série e que ofereça uma significativa redução na quantidade de informações quando comparado ao original. O valor máximo da janela é considerado o representativo do agrupamento.

Exemplificando, tem-se que com a janela de tamanho 5 e a série

{2; 7; 2; 5; 3; 8; 4; 4; 3; 2}

obtêm-se como contorno a nova série

{ 7; 8}

significando uma redução de 80% dos valores originais.

A seguir é apresentado um pseudocódigo ilustrativo sobre a operação do algoritmo ECJF considerando uma janela de tamanho 4.

```

carrega dados
normaliza dados

tam = tamanho(dados)
soma_dif = 0 // "soma_dif" significa soma das diferenças
janela = 4

nrContornos = arredonda(tam/janela)

para i de 1 até nrContornos
    aux = max(dados(:,(i-1)*janela+1:(i-1)*janela+janela))
    contorno(i) = aux
    soma_dif = soma_dif + soma(aux-dados(:,(i-1)*janela+1:(i-1)*janela+janela))
fim_para

saida = contorno

```

3.1.2 Janela variável

No método de Extração de Contornos por Janela Variável (ECJV) baseado no valor máximo de grupos, é realizada uma varredura ponto-a-ponto, e o valor máximo é considerado o representativo do agrupamento, com diferencial que o tamanho da janela não é fixo. Na sequência

$$\{2; 7; 2; 5; 3; 8; 4; 4; 3; 2\}$$

o algoritmo ECJV considera o maior valor do agrupamento, no caso, o valor

$$\{8\}$$

Contudo, este método pode gerar diversos resultados dependendo de qual seja o critério de erro admissível. Se o algoritmo estiver configurado, neste exemplo em análise, para apresentar erro de no máximo 12 unidades, teremos outra série representativa.

$$\{7; 8; 4\}$$

O que definirá um novo agrupamento são as diferenças acumuladas em valores absolutos. Quando a diferença acumulada for igual ou superior ao valor definido como critério de parada (erro admissível), a formação do agrupamento será encerrada, reiniciando o cálculo das diferenças acumuladas.

Tabela 2 - Disposição dos agrupamentos $a1$, $a2$ e $a3$ com uso janela variável e erros admissível máximo de 12 unidades

Janela	$a1$			$a2$			$a3$			
Posição	1	2	3	4	5	6	7	8	9	10
Valor	2	7	2	5	3	8	4	4	3	2
Diferenças acumuladas absolutas	-	5	10	3	5	10	4	4	5	6

Verifica-se na Tabela 2 que o agrupamento $a3$ ainda não foi encerrado, pois as diferenças acumuladas desse agrupamento não se igualaram ou ultrapassaram o limiar que, no exemplo exposto, é 12. Desta forma, o tamanho das janelas de agrupamento para o cenário acima apresentado foram, respectivamente, 3; 3 e 4.

A seguir é apresentado o pseudocódigo ilustrativo sobre a operação do algoritmo ECJV.

```

carrega dados
normaliza dados

criterio = 0.1
tam = tamanho(dados)
soma_dif = 0 // "soma_dif" significa soma das diferenças
diferencaTotal = 0

para i de 1 até tam-1

se soma_dif < criterio*diferenca_maxima
    indice_Min = i
    se dados(:,i+1) > dados(:,i)
        aux = dados(:,i+1)
        contorno(count) = aux
        indice_Max = i+1
        soma_dif = soma(aux-dados(:,indice_Min:indice_Max))
    fim_se

senao
    indice_Max = i+1
    diferencaTotal = diferencaTotal+soma_dif
    count = count+1
    soma_dif = 0
fim_senao

fim_para

saida = contornos

```

3.1.3 Critérios de escolha do método

Analisando os mecanismos de operação das janelas fixa e variável descritos nas subseções 3.1.1 e 3.1.2, observa-se que é necessário um parâmetro para limitar o tamanho da janela de agrupamento (JAG) dos algoritmos de extração de contornos. Para tal, foi inserido no algoritmo ECJV um critério de conclusão da JAG, o erro, a partir das diferenças acumuladas naquele agrupamento. Este critério, por sua vez, também pode ser ajustado por meio de uma constante de ajuste, denominada CAJ. Esta constante permite especificar qual o valor de erro é admissível na formação do agrupamento e, uma vez extrapolado este limite, encerra-se a JAG.

A Tabela 3 ilustra os processos para definição da JAG, tendo como parâmetro as diferenças acumuladas e redução. Entende-se por diferenças acumuladas as diferenças entre o valor máximo de um grupo, cujo tamanho será definido pela JAG, e o valor de

cada ponto da amostra (pacote); e de redução à capacidade do algoritmo em reduzir o volume de informações dos dados originais. Foram utilizados os traços da coleção CAIDA *Anonymized 2008 Internet Traces Dataset*⁴.

Tabela 3 - Determinação da JAG (traces CAIDA, 2000 pontos)

Tamanho JAG fixo (ECJF)			Tamanho JAG variável (ECJV)			
Tam. JAG	Diferenças acumuladas (diferença global) x 1000	Redução %	CAJ	Tam. JAG médio equivalente	Diferenças acumuladas (diferença global) x 1000	Redução %
2	0,9735	50,0000	0,0500	3,0488	0,400	67,2000
3	1,4536	66,7000	0,1000	3,2787	0,808	69,5000
4	1,7455	75,0000	0,1500	3,3445	1,232	70,1000
5	1,9379	80,0000	0,2000	3,3557	1,657	70,2000
6	2,0084	83,3500	0,2500	3,4483	2,089	71,0000
7	2,0713	85,7500	0,3000	3,4722	2,521	71,2000
8	2,0808	87,5000	0,3500	3,5587	2,955	71,9000
9	2,1273	88,9000	0,4000	3,5714	3,390	72,0000
10	2,1306	90,0000	0,4500	3,5971	3,825	72,2000
11	2,1164	90,9500	0,5000	3,6496	4,260	72,6000
12	2,1322	91,7000	0,5500	3,6900	4,696	72,9000
13	2,1383	92,3500	0,6000	3,8462	5,127	74,0000
14	2,1329	92,9000	0,6500	3,8911	5,559	74,3000
15	2,1487	93,3500	0,7000	3,8911	5,990	74,3000
16	2,1550	93,7500	0,7500	3,9683	6,419	74,8000
17	2,1472	94,1500	0,8000	3,9683	6,848	74,8000
18	2,1588	94,4500	0,8500	4,0161	7,274	75,1000
19	2,1543	94,7500	0,9000	4,0650	7,698	75,4000
20	2,1611	95,0000	0,9500	4,0984	8,120	75,6000
21	2,1564	95,2500	1,0000	4,1322	8,540	75,8000
22	2,1409	95,5000	1,0500	4,2017	8,956	76,2000
23	2,1402	95,7000	1,1000	4,2373	9,370	76,4000
24	2,1538	95,8500	1,1500	4,2735	9,781	76,6000
25	2,1652	96,0000	1,2000	4,3290	10,190	76,9000

De acordo com a Tabela 3, considerando os critérios Diferença Acumulada e Redução, a melhor relação se dá por meio da ECJV com CAJ = 0,100. Entretanto, a redução pode ser ainda maior dependendo da série de dados a ser analisada. Exemplificando, aplicamos o algoritmo ECJV com CAJ = 0,100 para a extração de contornos dos traços da Bellcore.

⁴ Traces disponíveis em http://www.caida.org/data/passive/passive_2008_dataset.xml

Neste conjunto de dados, observamos que foi obtida uma redução de 87,6 %, com diferenças acumuladas de $0,33 \times 10^3$ e tamanho da JAG médio de 8,06. Essa diferença se dá pelo grau de variabilidade entre os traços: quão menor for essa variabilidade, menos rapidamente o algoritmo ECJV irá encerrar um agrupamento, aumentando sua capacidade de redução.

Para os traços CAIDA foi utilizado JAG = 3, garantindo redução de 66,7% por meio do algoritmo ECJF.

3.2 Redes neurais aplicadas aos máximos de grupos

3.2.1 Justificativa

Motivados pelo fato do cérebro humano processar informações de forma diferente do computador digital, os estudos em redes neurais artificiais, ou simplesmente “redes neurais”, acarretaram em representativos avanços em reconhecimento de padrões, mineração de dados, monitoramento de tráfego e previsão de séries temporais. De forma mais genérica, diz-se que uma rede neural é projetada para modelar como o cérebro processa informações ao realizar uma determinada tarefa, utilizando-se, para isso, de circuitos ou componentes eletrônicos, ou simplesmente simulando por programação em um computador digital (Haykin, 1994).

São vários os motivos que culminou a aplicação de redes neurais artificiais na previsão de tráfego IP; entre eles podemos citar (i) a sua capacidade de mapeamento entrada-saída não-lineares; (ii) o seu aprendizado por exemplos; (iii) a capacidade de adaptação a novas situações pela mudança dos pesos; e (iv) a sua capacidade de decisão diante de dados ruidosos ou mesmo corrompidos. Para este trabalho, um fator determinante para experimentações utilizando as redes neurais é a necessidade de modelos preditores acurados na estimativa de traços de rede.

De forma geral, redes neurais artificiais podem ser divididas quanto ao tipo de aprendizado em duas categorias: (i) redes com aprendizado supervisionado e (ii) redes com aprendizado não-supervisionado. No caso supervisionado, cada entrada apresentada à rede vem acompanhada de uma saída desejada, a fim de permitir uma modificação dos parâmetros ajustáveis em função do erro entre a resposta fornecida pela rede e a saída real desejada. Ao final da etapa de ajuste dos parâmetros, chamada genericamente de

treinamento, as respostas da rede para todas as entradas devem ser próximas das saídas desejadas. No caso não-supervisionado, a rede neural detecta padrões e características estatísticas do espaço de entrada, de forma a construir uma representação de dimensionalidade reduzida do mesmo no conjunto de pesos sinápticos de seus neurônios (Menezes Jr., 2006).

3.2.2 Rede Perceptron Multicamadas

Foram utilizadas as arquiteturas neurais *MultiLayer Perceptron* (MLP) e *Nonlinear AutoRegressive with exogenous input* (NARX), ambas com aprendizado online, sem etapas distintas de treinamento e execução. A Figura 2 mostra a arquitetura de uma rede neural multicamadas com apenas uma camada escondida.

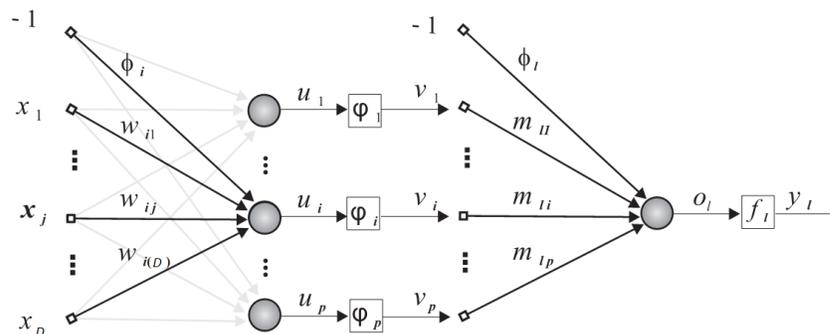


Figura 2 - Arquitetura de uma MLP com uma camada escondida (adaptado de Aragão Jr., 2008)

O número de neurônios na camada de saída e o tamanho do vetor de entrada do módulo neural proposto foram definidos de acordo com a estrutura do problema atacado nesta dissertação. A variável de saída é o valor do contorno a ser estimado, ou seja, a rede neural apresenta um neurônio na camada de saída.

Uma janela com tamanho muito grande pode fazer com que a rede responda lentamente a mudanças bruscas no atraso enquanto janelas muito pequenas podem não conter a quantidade de informação necessária para uma boa previsão (Aragão Jr., 2008).

3.2.2.1 Dimensão do vetor de entrada (D)

Considerando a matriz M representativa dos contornos dos traços de tráfego, temos que:

$$D = m \times n + 1, \quad (12)$$

em que m representa a quantidade de amostras do contorno e n a informação representativa, a saber: tamanho representativo normalizado do pacote. Empiricamente, verificou-se ser suficiente cinco amostras para a rede neural inferir o próximo elemento. Logo a dimensão do vetor de entrada (D) é determinado como

$$D = 5 \times 1 + 1,$$

$$D = 6.$$

3.2.2.2 Determinação do número de neurônios ocultos

Normalmente, utilizam-se modelos conexionistas com uma ou duas camadas escondidas. Os modelos com uma camada escondida são capazes de aproximar com precisão arbitrária (dentro de um limite de erro) funções contínuas, enquanto os de duas ou mais camadas escondidas podem fazer aproximação inclusive de funções descontínuas (Aragão Jr., 2008). Fazendo-se uso do Princípio da Navalha de Occam⁵, os testes têm início com uma camada escondida, e apenas se os resultados não forem satisfatórios deve-se incrementar a quantidade de camadas escondidas.

O número de neurônios da camada escondida define a complexidade e o poder computacional da rede neural. Apesar de existir algumas fórmulas heurísticas para encontrar esses valores, a prática corrente é a empírica, por tentativa e erro, analisando principalmente o erro quadrático médio (MSE) gerado pela saída. Adicionalmente, os resultados neste trabalho foram analisados por meio da Métrica das Amplitudes Conjugadas (MAC). Detalhes da métrica MAC estão presentes na Seção 4.2.

Para determinar a quantidade de neurônios na camada escondida na rede MLP, foram realizados experimentos com 52, 54, 56, ..., 96, 98 e 100 neurônios, sendo que o melhor resultado foi obtido com 54 neurônios na camada oculta (vide Tabela 4).

As configurações para os traços originais e contornos por ECJF e ECJV seguem procedimentos similares, por isso será apresentada detalhadamente apenas a metodologia para os contornos obtidos pelo algoritmo ECJV. As configurações para os demais traços serão apresentadas na conclusão desta seção.

A Tabela 4 ilustra a metodologia de escolha do número de neurônios na camada escondida considerando o MSE e MAC médios de dez rodadas de simulação para cada

⁵ Também chamado de *princípio da pluralidade desnecessária*, é um princípio lógico atribuído ao frade Franciscano inglês William de Ockham. O princípio recomenda que se escolha a teoria explicativa que implique o menor número de premissas assumidas e o menor número de entidades.

uma das configurações de neurônios ocultos; e o tempo total, em segundos, das dez rodadas. Nesta etapa foram fixadas arbitrariamente a constante de aprendizagem $ETA = 0,01$ e o valor de momentum $m = 0$, para que fossem ajustadas em seguida. Os testes foram executados com o uso dos contornos obtidos pelo algoritmo ECJV. O critério de escolha do número de neurônios é composto por uma relação ponderada descrita através da equação:

$$C_{rp} = \frac{(tempo_{TT} + 2 \times MAC_m + 3 \times MSE_m)}{6}, \quad (13)$$

onde C_{rp} é o critério com relação ponderada, $tempo_{TT}$ é o tempo total em segundos das 10 rodadas, MAC_m é o MAC médio das 10 rodadas e MSE_m é o MSE médio das 10 rodadas. O menor valor de C_{rp} indica, seguindo a relação proposta, a quantidade de neurônios a ser eleita. Na Tabela 4 observa-se que o menor valor da relação ponderada foi 0,1513, indicando que a quantidade de neurônios ocultos deve ser 54. Os traços utilizados são da coleção CAIDA.

Tabela 4 - Determinação da qtde. de neurônios ocultos na rede MLP ($ETA = 0,01$ e $m = 0$)

Nr. de neurônios	MSE médio das 10 rodadas	MAC médio das 10 rodadas	Tempo (s) total das 10 rodadas	Critério (relação ponderada)
52	0,0817	0,0491	0,6216	0,1608
54	0,0810	0,0478	0,5694	0,1513
56	0,0808	0,0462	0,5782	0,1522
58	0,0806	0,0446	0,5885	0,1533
60	0,0804	0,0430	0,6163	0,1572
62	0,0803	0,0417	0,6357	0,1600
64	0,0801	0,0405	0,6072	0,1548
66	0,0800	0,0391	0,6135	0,1553
68	0,0799	0,0380	0,6047	0,1534
70	0,0798	0,0370	0,6324	0,1576
72	0,0796	0,0359	0,6568	0,1612
74	0,0795	0,0349	0,6623	0,1618
76	0,0795	0,0341	0,6710	0,1629
78	0,0794	0,0332	0,6806	0,1642
80	0,0792	0,0321	0,6854	0,1646
82	0,0791	0,0313	0,7071	0,1678
84	0,0790	0,0304	0,7341	0,1720
86	0,0790	0,0298	0,8011	0,1829
88	0,0789	0,0290	0,7473	0,1737
90	0,0789	0,0285	0,7560	0,1749
92	0,0788	0,0276	0,7772	0,1781
94	0,0787	0,0270	0,7993	0,1816
96	0,0787	0,0267	0,8252	0,1858
98	0,0786	0,0258	0,8125	0,1833
100	0,0786	0,0255	0,8369	0,1873

A Figura 3 ilustra a projeção do MSE e MAC médios e o Tempo total para a rede MLP a fim de se obter um valor adequado para o número de neurônios na camada escondida.

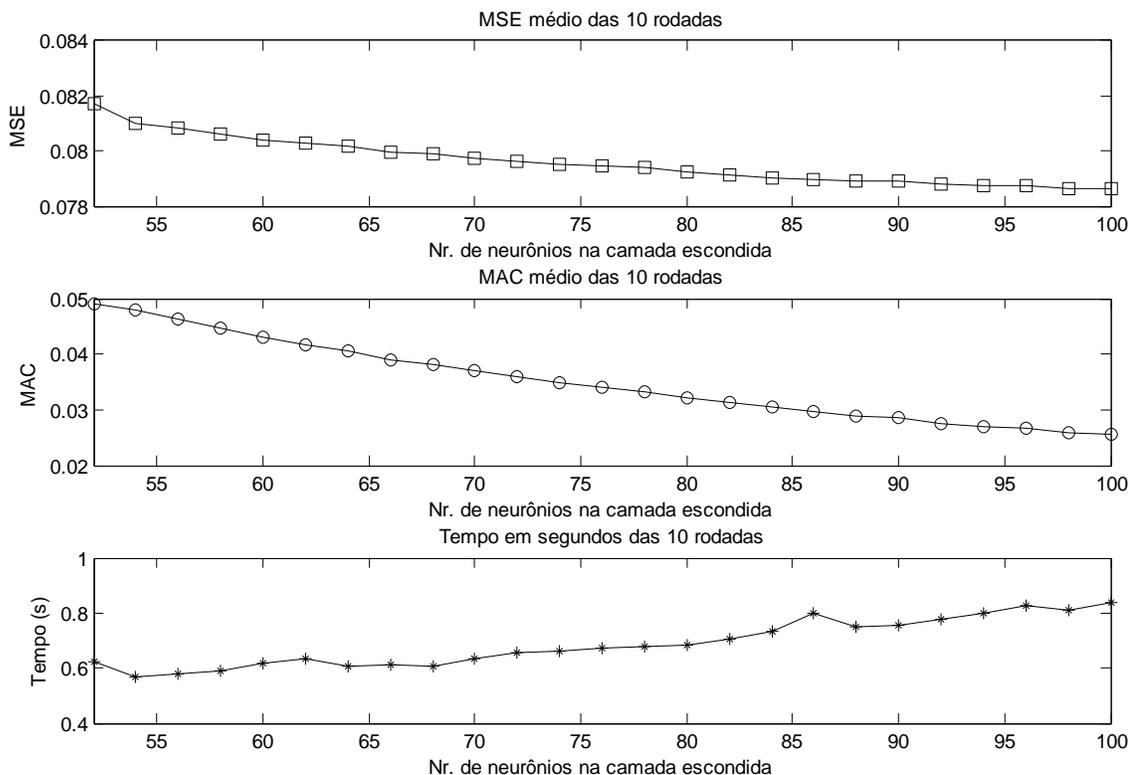


Figura 3 - Curvas das Figuras de mérito usadas na determinação do número de neurônios na rede MLP

3.2.2.3 Taxa de aprendizagem

De posse da estrutura da rede neural até o momento definida (54 neurônios ocultos e $m = 0$), os testes foram executados com o uso dos contornos obtidos pelo algoritmo ECJV e tiveram continuidade processando-se a taxa de aprendizagem (ETA), obtendo-se os valores expressos na Tabela 5.

Fazendo uso do critério com relação ponderada Crp adota-se a taxa de aprendizagem de valor 0,046. O menor valor de Crp indica, seguindo a relação proposta, a taxa de aprendizagem a ser eleita. Verifica-se ainda que se adotando $ETA = 0,046$ obtém-se um valor MAC próximo a zero, mantendo MSE e Tempo baixos. A função de ativação utilizada foi sigmóide logística.

Tabela 5 - MSE e MAC da simulação neural com MLP, alternando-se a taxa de aprendizagem

Taxa de aprendizagem (ETA)	MSE médio das 10 rodadas	MAC médio das 10 rodadas	Tempo (s) total das 10 rodadas	Critério (relação ponderada)
0,0020	0,1003	0,1716	0,4375	0,1803
0,0040	0,0886	0,1077	0,4049	0,1477
0,0060	0,0844	0,0764	0,3771	0,1305
0,0080	0,0823	0,0588	0,3745	0,1232
0,0100	0,0811	0,0478	0,3789	0,1196
0,0120	0,0803	0,0402	0,3713	0,1154
0,0140	0,0798	0,0347	0,3826	0,1152
0,0160	0,0795	0,0306	0,3831	0,1138
0,0180	0,0792	0,0274	0,4047	0,1162
0,0200	0,0791	0,0248	0,3756	0,1104
0,0220	0,0790	0,0227	0,3770	0,1099
0,0240	0,0789	0,0209	0,3909	0,1116
0,0260	0,0789	0,0195	0,3761	0,1086
0,0280	0,0789	0,0181	0,3777	0,1084
0,0300	0,0789	0,0171	0,3759	0,1078
0,0320	0,0789	0,0160	0,4115	0,1134
0,0340	0,0790	0,0151	0,3822	0,1082
0,0360	0,0790	0,0144	0,4068	0,1121
0,0380	0,0791	0,0137	0,3897	0,1091
0,0400	0,0792	0,0131	0,3711	0,1058
0,0420	0,0793	0,0125	0,3779	0,1068
0,0440	0,0794	0,0120	0,3872	0,1082
0,0460	0,0795	0,0115	0,3679	0,1049
0,0480	0,0796	0,0111	0,3768	0,1063
0,0500	0,0797	0,0107	0,3720	0,1054
0,0520	0,0799	0,0103	0,3760	0,1061
0,0540	0,0800	0,0100	0,3903	0,1084
0,0560	0,0801	0,0097	0,3763	0,1060
0,0580	0,0803	0,0094	0,3777	0,1062
0,0600	0,0804	0,0091	0,3898	0,1082

A Figura 4 ilustra a projeção do MSE e MAC médios e o Tempo total para a rede MLP a fim de se obter um valor adequado para a taxa de aprendizagem *ETA*.

3.2.2.4 Valor do termo de momento

Nesta etapa objetiva-se averiguar o valor do termo de momento que ofereça valores de MAC e MSE próximos a zero e baixo tempo de execução. Alternando-se o valor deste parâmetro na rede MLP, obtêm-se os valores expressos na Tabela 6.

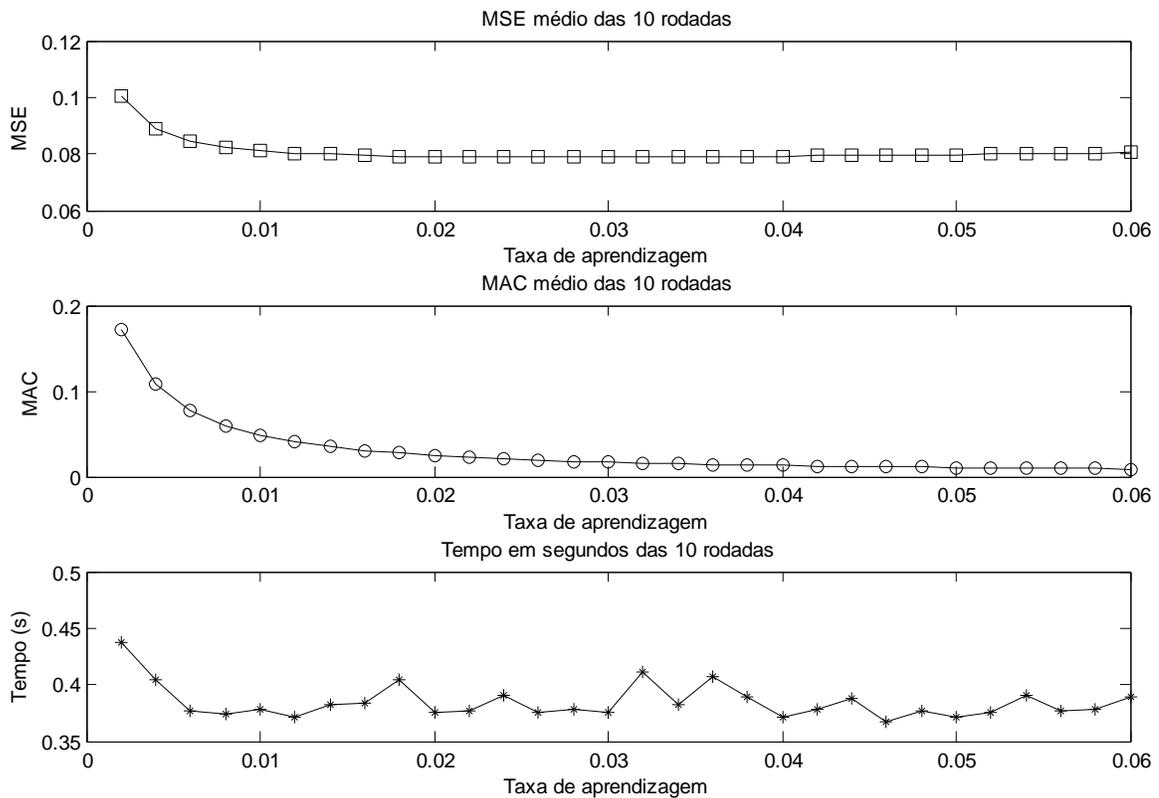


Figura 4 - Curvas das Figuras de mérito usadas na determinação da taxa de aprendizagem na rede MLP

Tabela 6 - MSE da simulação neural com MLP, alternando-se o valor do termo de momento

Termo de momento	MSE médio das 10 rodadas	MAC médio das 10 rodadas	Tempo (s) total das 10 rodadas	Critério (relação ponderada)
0	0,0796	0,0106	0,5128	0,1288
0,05	0,0793	0,0110	0,4596	0,1199
0,10	0,0795	0,0105	0,4530	0,1188
0,15	0,0797	0,0100	0,4378	0,1161
0,20	0,0799	0,0094	0,4429	0,1169
0,25	0,0801	0,0090	0,4346	0,1155
0,30	0,0804	0,0084	0,4547	0,1188
0,35	0,0808	0,0079	0,4346	0,1155
0,40	0,0813	0,0074	0,4625	0,1202
0,45	0,0818	0,0069	0,4374	0,1161
0,50	0,0825	0,0064	0,4435	0,1173
0,55	0,0833	0,0059	0,4377	0,1166
0,60	0,0844	0,0055	0,4595	0,1206

O valor do termo de momento adotado foi **0,35**, visto que o critério de relação ponderada *Crp* apresentou menor valor nesta configuração. A Figura 5 ilustra a projeção do MSE e MAC médios e o Tempo total para a rede MLP a fim de se obter um valor adequado para o termo de momento. Os testes foram executados com o uso dos contornos obtidos pelo algoritmo ECJV.

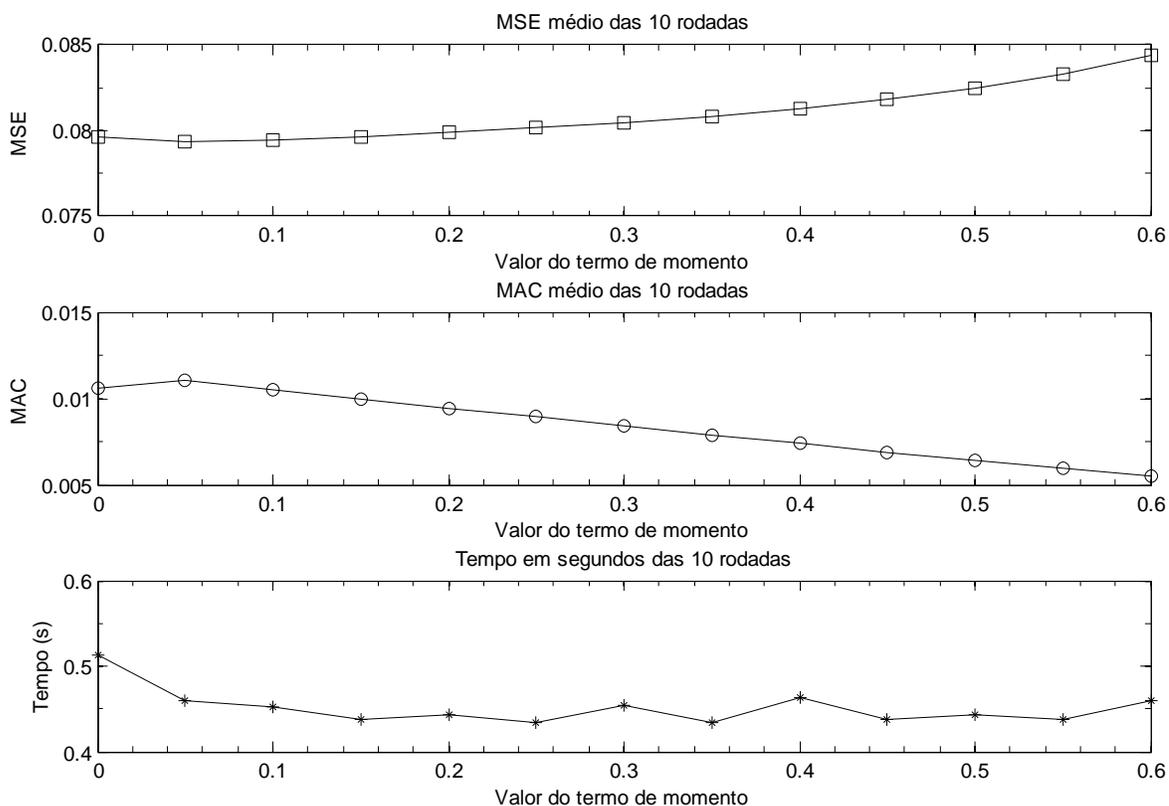


Figura 5 - Curvas das Figuras de mérito usadas na determinação do termo de momento na rede MLP

3.2.6 Topologia NARX

Para efeito comparativo da proposta, adotou-se uma segunda arquitetura neural. Uma importante e útil classe de sistemas não-lineares de tempo discreto é matematicamente representada pelo modelo NARX (*Nonlinear AutoRegressive model with eXogenous inputs*) (Leontaritis & Billings, 1985) e (Norgaard *et al.*, 2000).

O modelo NARX é aplicado para diversos tipos de séries temporais com padrões caóticos ou fractais aplicadas como entrada para a rede neural (Diaconescu, 2008). A Figura 6 apresenta uma das possíveis arquitetura da rede NARX.

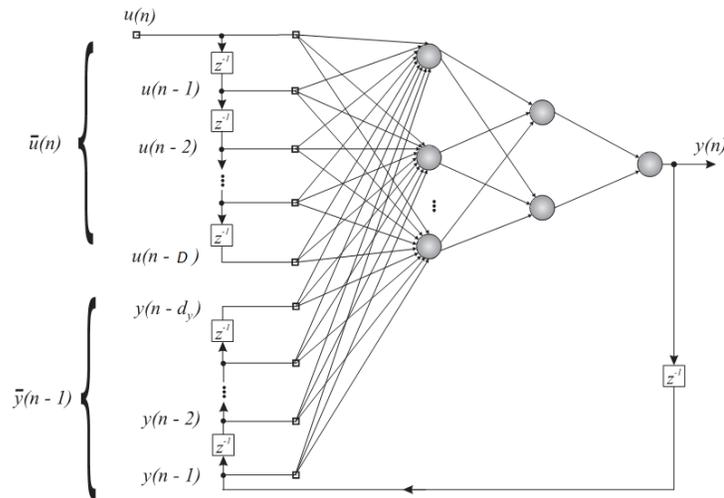


Figura 6 - Arquitetura de uma NARX com D entradas (adaptado de Menezes Jr., 2006)

3.2.6.1 Dimensão do vetor de entrada D

De forma semelhante à rede MLP, considerando a matriz M representativa dos contornos dos traços de tráfego, tem-se que

$$D = m \times n + 1 + r \quad (14)$$

em que m representa cinco amostras do contorno, n a informação representativa e r as entradas de retroalimentação; logo, a dimensão do vetor de entrada (D) é dada por

$$D = 5 \times 1 + 1 + 5,$$

$$D = 11$$

3.2.6.2 Determinação do número de neurônios ocultos

Para as redes NARX, foram realizados experimentos com 52, 54, 56, ..., 96, 98 e 100 neurônios e o melhor resultado foi obtido com o uso de 54 neurônios na camada oculta, (vide Tabela 7), quando verifica-se que o menor valor do critério de relação ponderada Crp , descrita na Equação 14. De forma semelhante à configuração da rede MLP, fazendo-se uso do Princípio da Navalha de Occam, os testes têm início com uma camada escondida, e apenas se os resultados não forem satisfatórios deve-se incrementar a quantidade de camadas escondidas.

As configurações para os traços originais e contornos por ECJF e ECJV seguem procedimentos similares, por isso será apresentada detalhadamente apenas a metodologia para os contornos obtidos pelo algoritmo ECJV. As configurações para os demais traços serão apresentadas na conclusão desta seção.

A Tabela 7 ilustra a metodologia de escolha do número de neurônios na camada escondida considerando o MSE e MAC médios de dez rodadas de simulação para cada uma das configurações de neurônios ocultos; e o tempo total, em segundos, das dez rodadas.

Nesta etapa foram fixadas arbitrariamente a constante de aprendizagem $ETA = 0,01$ e o valor de momentum $m = 0$, para que fossem ajustadas em seguida. Os traços utilizados são da coleção CAIDA.

Tabela 7 - Determinação da qtde. de neurônios ocultos na rede NARX ($ETA = 0,01$ e $m = 0$)

Nr. de neurônios	MSE médio das 10 rodadas	MAC médio das 10 rodadas	Tempo (s) total das 10 rodadas	Critério (relação ponderada)
52	0,0818	0,0457	0,5445	0,1469
54	0,0811	0,0442	0,4752	0,1344
56	0,0809	0,0428	0,4936	0,1370
58	0,0807	0,0412	0,5284	0,1422
60	0,0805	0,0396	0,5534	0,1457
62	0,0804	0,0385	0,5404	0,1431
64	0,0802	0,0369	0,5614	0,1460
66	0,0801	0,0357	0,5746	0,1477
68	0,0800	0,0346	0,5901	0,1499
70	0,0798	0,0333	0,6110	0,1528
72	0,0797	0,0325	0,6472	0,1585
74	0,0796	0,0315	0,6841	0,1643
76	0,0795	0,0304	0,6206	0,1533
78	0,0794	0,0296	0,6460	0,1572
80	0,0793	0,0288	0,6595	0,1592
82	0,0793	0,0280	0,6722	0,1610
84	0,0791	0,0270	0,7211	0,1688
86	0,0790	0,0262	0,7335	0,1705
88	0,0790	0,0257	0,7392	0,1712
90	0,0789	0,0250	0,7530	0,1733
92	0,0789	0,0242	0,8054	0,1817
94	0,0788	0,0235	0,7995	0,1805
96	0,0788	0,0231	0,8307	0,1855
98	0,0787	0,0225	0,8254	0,1844
100	0,0786	0,0219	0,8433	0,1871

A Figura 7 ilustra a projeção do MSE e MAC médios e o Tempo total para a rede NARX a fim de se obter um valor adequado para o número de neurônios na camada escondida.

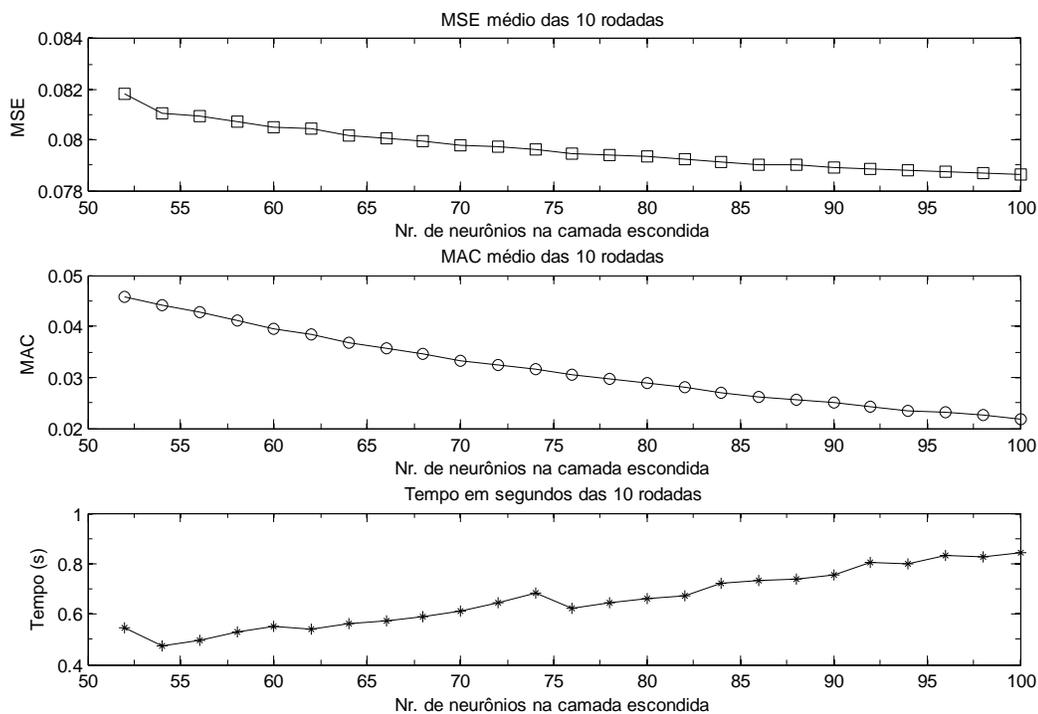


Figura 7 - Curvas das Figuras de mérito usadas na determinação do número de neurônios na camada escondida da rede NARX

3.2.6.3 Constante de aprendizagem

De posse da estrutura da rede neural escolhida (54 neurônios ocultos e $m = 0$), os testes tiveram continuidade processando-se a taxa de aprendizagem e obtendo-se os valores expressos na Tabela 8.

Fazendo uso do critério com relação ponderada Crp adota-se a taxa de aprendizagem de valor 0,05. O menor valor de Crp indica, seguindo a relação proposta, a taxa de aprendizagem a ser eleita. Verifica-se ainda que se adotando $ETA = 0,05$ obtém-se um valor MAC próximo a zero, mantendo MSE e Tempo baixos. A função de ativação utilizada foi sigmóide logística.

Tabela 8 - MSE e MAC da simulação neural com NARX, alternando-se a constante de aprendizagem

Taxa de aprendizagem (ETA)	MSE médio das 10 rodadas	MAC médio das 10 rodadas	Tempo (s) total das 10 rodadas	Critério (relação ponderada)
0,0020	0,1005	0,1662	0,5533	0,1979
0,0040	0,0888	0,1033	0,5273	0,1667
0,0060	0,0845	0,0727	0,5343	0,1555
0,0080	0,0824	0,0553	0,5500	0,1513
0,0100	0,0812	0,0444	0,5197	0,1420
0,0120	0,0804	0,0367	0,5353	0,1417
0,0140	0,0799	0,0313	0,5435	0,1410
0,0160	0,0795	0,0270	0,5498	0,1404
0,0180	0,0793	0,0238	0,5423	0,1379
0,0200	0,0791	0,0212	0,5495	0,1382
0,0220	0,0790	0,0192	0,5431	0,1364
0,0240	0,0789	0,0173	0,5372	0,1347
0,0260	0,0789	0,0158	0,5510	0,1365
0,0280	0,0789	0,0146	0,5309	0,1328
0,0300	0,0789	0,0134	0,5365	0,1333
0,0320	0,0789	0,0125	0,5364	0,1330
0,0340	0,0790	0,0116	0,5368	0,1328
0,0360	0,0790	0,0108	0,5263	0,1308
0,0380	0,0791	0,0101	0,5408	0,1330
0,0400	0,0792	0,0095	0,5417	0,1330
0,0420	0,0793	0,0089	0,5288	0,1307
0,0440	0,0793	0,0083	0,5319	0,1311
0,0460	0,0795	0,0079	0,5288	0,1305
0,0480	0,0796	0,0074	0,5311	0,1308
0,0500	0,0797	0,0071	0,5258	0,1298
0,0520	0,0798	0,0067	0,5398	0,1321
0,0540	0,0799	0,0063	0,5284	0,1301
0,0560	0,0801	0,0060	0,5293	0,1303
0,0580	0,0802	0,0057	0,5354	0,1313
0,0600	0,0804	0,0054	0,5355	0,1312

A Figura 8 ilustra a projeção do MSE e MAC médios e o Tempo total para a rede MLP a fim de se obter um valor adequado para a constante de aprendizagem *ETA*.

3.2.6.4 Termo de momento

Nesta etapa objetiva-se averiguar o valor do termo de momento que ofereça valores de MAC e MSE próximos a zero e baixo tempo de execução. Alternando-se o valor deste parâmetro na rede NARX, obtêm-se os valores expressos na Tabela 9.

O valor do termo de momento adotado foi 0,15, visto que o critério de relação ponderada Crp apresentou menor valor nesta configuração. A Figura 9 ilustra a projeção do MSE e MAC médios e o Tempo total para a rede NARX a fim de se obter um valor adequado para o termo de momento.

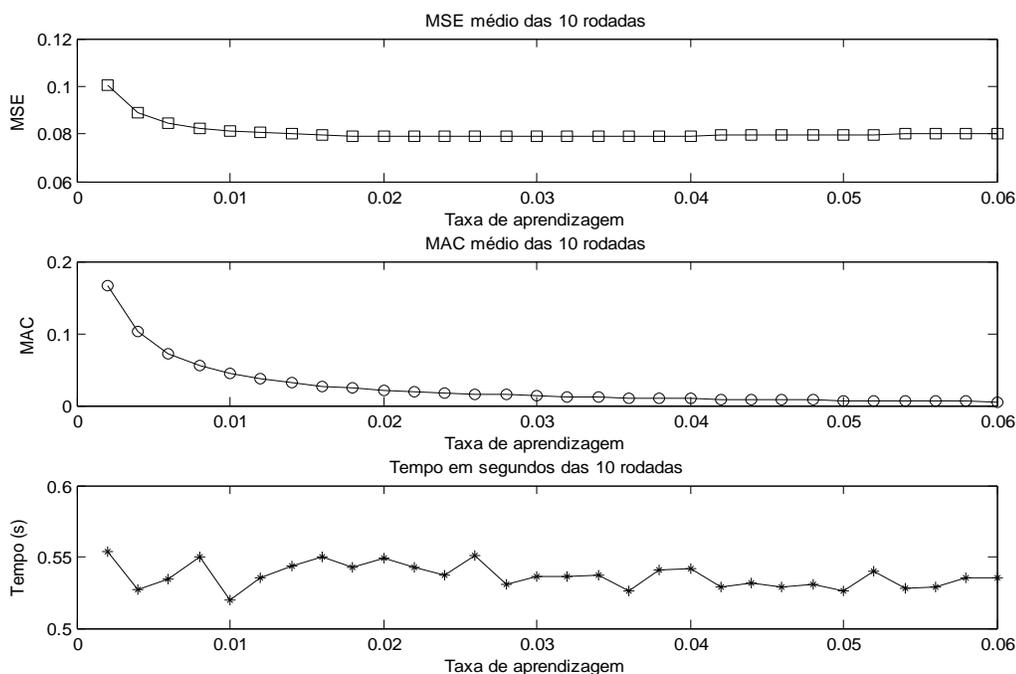


Figura 8 - Curvas das Figuras de mérito usadas na determinação da constante de aprendizagem na rede NARX

Tabela 9 - MSE da simulação neural com NARX, alternando-se o valor de momentum

Termo de momento	MSE médio das 10 rodadas	MAC médio das 10 rodadas	Tempo (s) total das 10 rodadas	Critério (relação ponderada)
0	0,0798	0,0061	0,6082	0,1433
0,05	0,0795	0,0066	0,5761	0,1380
0,10	0,0797	0,0061	0,5178	0,1282
0,15	0,0799	0,0056	0,4948	0,1243
0,20	0,0802	0,0051	0,5179	0,1281
0,25	0,0805	0,0046	0,5087	0,1266
0,30	0,0808	0,0042	0,4996	0,1251
0,35	0,0812	0,0037	0,5109	0,1270
0,40	0,0818	0,0032	0,5048	0,1261
0,45	0,0824	0,0027	0,5089	0,1269
0,50	0,0832	0,0022	0,4979	0,1253
0,55	0,0842	0,0018	0,5141	0,1284
0,60	0,0855	0,0014	0,5320	0,1319

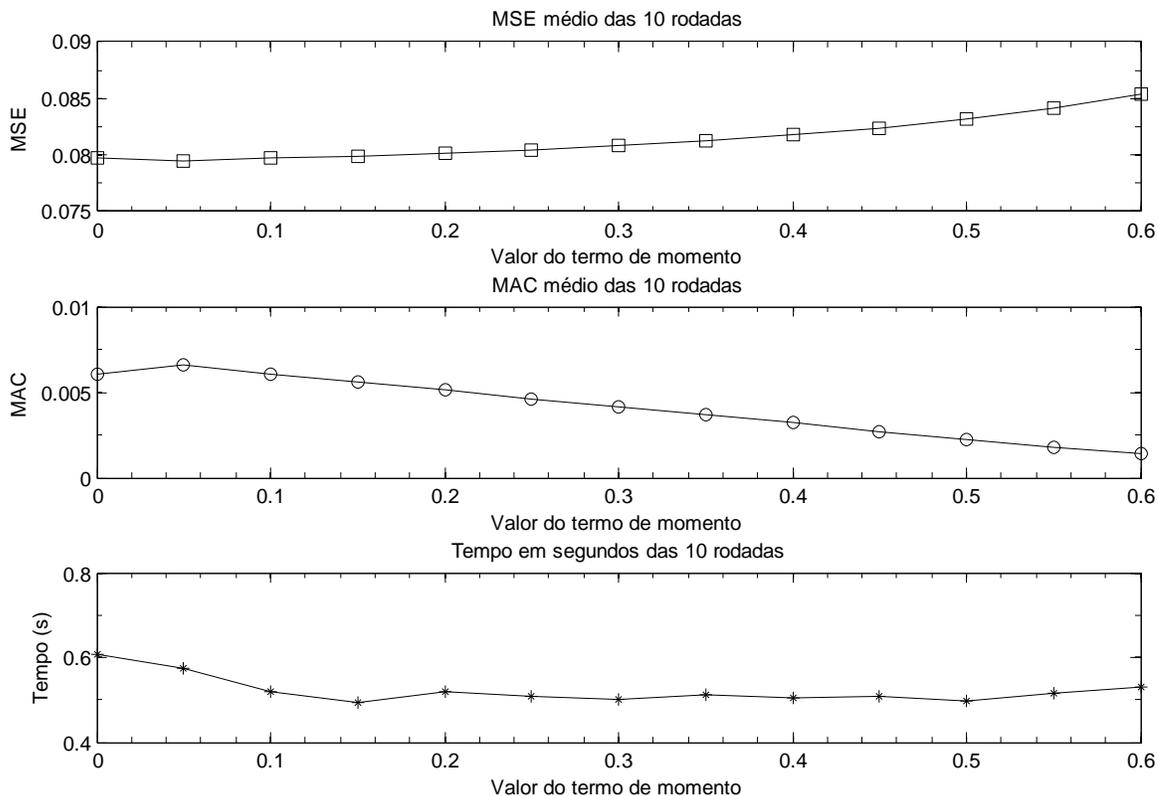


Figura 9 - Curvas das Figuras de mérito usadas na determinação do termo de momento na rede NARX

3.3 Conclusão

A redução no número de amostras a serem lançadas a entrada da rede neural foi obtida com o uso do algoritmo Extrator de Contornos. O algoritmo proposto realiza uma extração dos contornos de uma série temporal, chegando a reduzir em até 95% a informação original (tamanho da sequência).

Para o algoritmo ECJF foi adotado $JAG = 3$, garantindo uma redução de 66,7%; para o ECJV foi adotada a $CAJ = 0,1$, resultando em uma JAG equivalente de 3,28 com redução aproximada de 69,5%.

O resumo das configurações neurais da rede MLP para os traços originais e contornos segue na Tabela 10.

Tabela 10 - Resumo da configuração da rede MLP

	Traços originais	Contornos por ECJF	Contornos por ECJV
Nr. Neurônios ocultos	54	56	54
Taxa de aprendizagem	0,006	0,034	0,046
Valor de momento	0,15	0,25	0,35

O resumo das configurações neurais da rede NARX para os traços originais e contornos é apresentado na Tabela 11.

Tabela 11 - Resumo da configuração da rede NARX

	Traços originais	Contornos por ECJF	Contornos por ECJV
Nr. Neurônios ocultos	54	56	54
Taxa de aprendizagem	0,008	0,026	0,05
Valor de momento	0,05	0,35	0,15

Capítulo 4

Resultados e discussão

No decorrer das pesquisas foram utilizados traços de tráfego da coleção CAIDA *Anonymized 2008 Internet Traces Dataset*, da rede *Equinix* em Chicago, datados de 2008. Os dados foram coletados de um enlace bidirecional OC-192 (Optical Carrier, 10 GBps), de backbone Internet. Devido ao formato “pcap” dos arquivos dos traços, utilizou-se o aplicativo Wireshark⁶ versão 1.2.3 para se obter os tamanhos dos pacotes contidos na coleção. Foram utilizadas também as ferramentas de computação Minitab⁷ versão 15.1 e o ambiente computacional Matlab⁸ versão 7.6.

4.1 Caracterização e metodologia da simulação

Os dados originais precisam ser processados pelo algoritmo ECJV para a obtenção dos contornos dos traces. Tal procedimento se faz necessário para reduzir o número de informações a serem lançadas à entrada do mecanismo preditor. De posse de um novo conjunto de dados, este é submetido aos mecanismos estatísticos MA, AES e HW; e neurocomputacionais, MLP e Narx. Os resultados obtidos são confrontados estatisticamente e por meio da Métrica de Amplitudes Conjugadas fornecendo indicações da capacidade de redução dos contornos, e de um modelo para sua predição. A Figura 10 apresenta uma ilustração da metodologia de avaliação adotada. Em síntese, os dados originais são processados com o algoritmo extrator de contornos e em seguida são submetidos ao métodos de predição, constituindo 18 cenários. Em cada cenário são

⁶ Aplicativo Wireshark disponível em <http://www.wireshark.org/>

⁷ Aplicativo Minitab disponível em <http://www.minitab.com/>

⁸ Aplicativo Matlab disponível em <http://www.mathworks.com/>

realizadas 10 rodadas de simulações. Em seguida, os conjuntos de dados obtidos são analisados pelas métricas MSE e MAC.

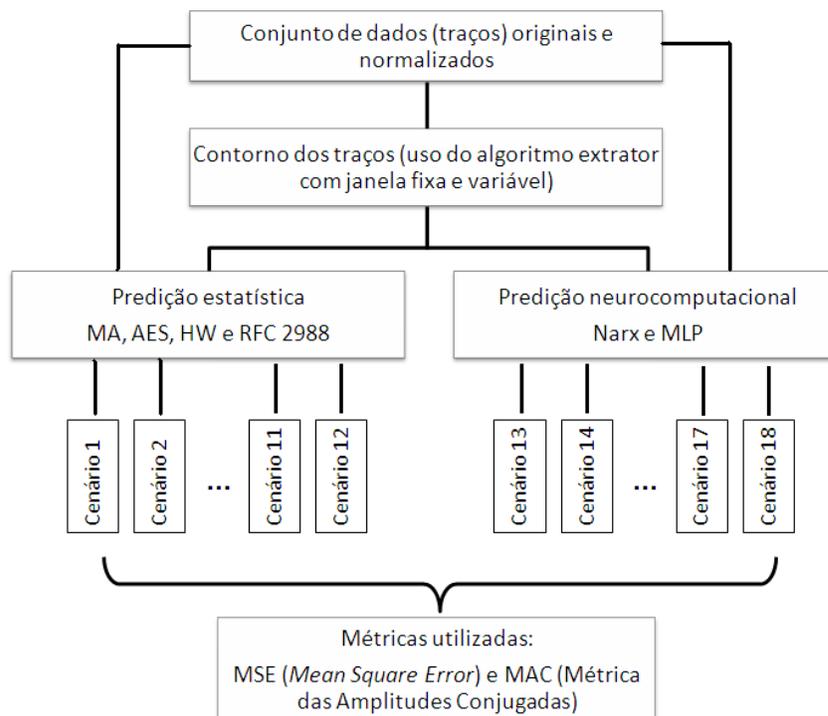


Figura 10 - Metodologia aplicada

Realizando as relações possíveis entre os 6 métodos em análise e os 3 tipos de dados disponíveis (traço original, contornos por ECJF e por ECJV), chega-se a 18 cenários de simulações, conforme ilustrado na Tabela 12.

Tabela 12 - Cenários de teste

	Traços originais	Contornos por ECJF	Contornos por ECJV
MA	Cenário 1	Cenário 2	Cenário 3
AES	Cenário 4	Cenário 5	Cenário 6
HW	Cenário 7	Cenário 8	Cenário 9
RFC 2988	Cenário 10	Cenário 11	Cenário 12
MLP	Cenário 13	Cenário 14	Cenário 15
Narx	Cenário 16	Cenário 17	Cenário 18

Para cada um dos 18 cenários, foram executadas 10 rodadas de simulações utilizando-se trechos adjacentes, sem sobreposições, do traço Caida original.

4.2 Métricas analisadas

A fim de se considerar a quantidade de acertos do tamanho de pacotes (ou grupos) globalmente, independente da incidência de erros pontuais, foi desenvolvida a Métrica das Amplitudes Conjugadas (MAC), que realiza uma comparação indireta entre os tamanhos médios dos agrupamentos ou pacotes originais e dos preditos.

Diferentemente da métrica MSE, MAC não considera o aspecto temporal, ou seja, não é influenciada por uma predição defasada ou adiantada. A métrica MAC retorna valores próximos ao *zero*, de forma que quanto mais próximo de *zero* for o resultado, melhor o estimador com relação à capacidade de dimensionar a vazão necessária à rede.

A Equação abaixo representa a métrica MAC:

$$MAC = 1 - \frac{\sum_{i=1}^{size(sP)} sP(i)}{size(sP)} \times \frac{size(sO)}{\sum_{i=1}^{size(sO)} sO(i)}, \quad (15)$$

em que sP representa a série predita pelos módulos preditores, sO representa a série original a ser comparada, $size$ indica a função que extrai o tamanho da série, abs indica a função que retorna o valor absoluto da amplitude, e i indica cada um dos pontos da série em questão. Da Equação 16 pode-se obter sua forma reduzida em

$$MAC = 1 - \frac{\sum_{i=1}^{size(sP)} sP(i)}{\sum_{i=1}^{size(sO)} sO(i)}, \quad (16)$$

A métrica MAC possui como característica positiva permanecer imune a alterações quando os dados estão dispostos normalizados.

Além da métrica MAC, é utilizada a métrica MSE como meio de quantificar a diferença entre o valor estimado e o original.

4.3 Resultados

Nesta seção serão apresentados e discutidos os 18 cenários que compõem esta dissertação.

4.3.1 Cenário 1: MA com dados originais

Neste cenário aplica-se o método MA aos dados originais normalizados. A Figura 11 ilustra a melhor das dez simulações de predição entre a série original e a predita. É apresentado o resultado da predição. A Figura 12 apresenta o histograma do erro para a simulação. As análises foram efetuadas com MA centrado em 4 termos.

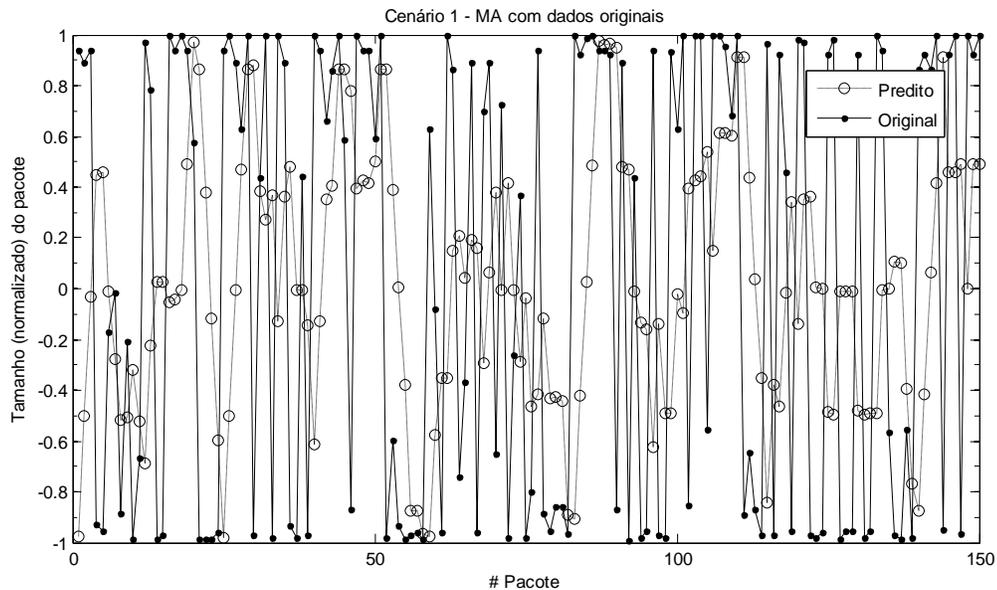


Figura 11 - Cenário 1 - Método MA com os dados originais normalizados

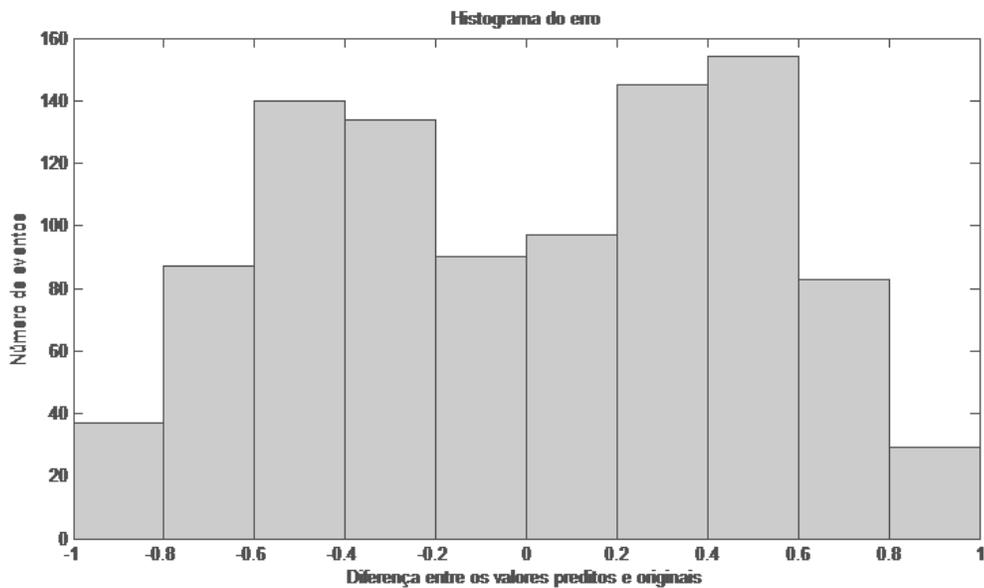


Figura 12 - Cenário 1 - Histograma do erro

Quanto mais centrado em torno de *zero* estiver o histograma do erro melhor será o resultado do modelo preditor. No Cenário 1 observamos que os valores preditos não obtiveram exatidão, resultando em um histograma do erro disperso. Isto indica que o método MA associado a dados originais não é uma abordagem segura para a obtenção de uma predição precisa.

4.3.2 Cenário 2: MA com contornos por ECJF

Este cenário envolve a aplicação do método MA à série de contorno obtida pelo algoritmo ECJF. A Figura 10 ilustra a melhor relação entre a série original e a predita, entre 10 simulações realizadas.

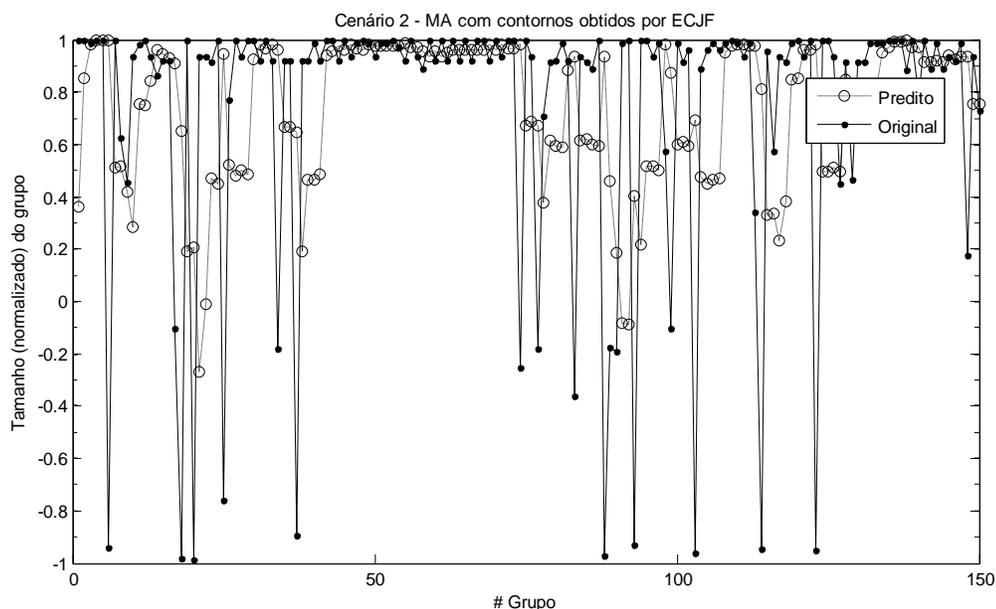


Figura 13 - Cenário 2 - Método MA com os contornos obtidos por ECJF

A partir da Figura 14 observa-se pelo histograma do erro que este está concentrado em um conjunto de valores próximo a 0,3, indicando que houve erro recorrente na predição. Isto indica que o método MA associado a contornos obtidos através do algoritmo ECJF não é uma boa abordagem para a obtenção de uma predição precisa, contudo, mostra-se melhor do que se utilizando de dados originais.

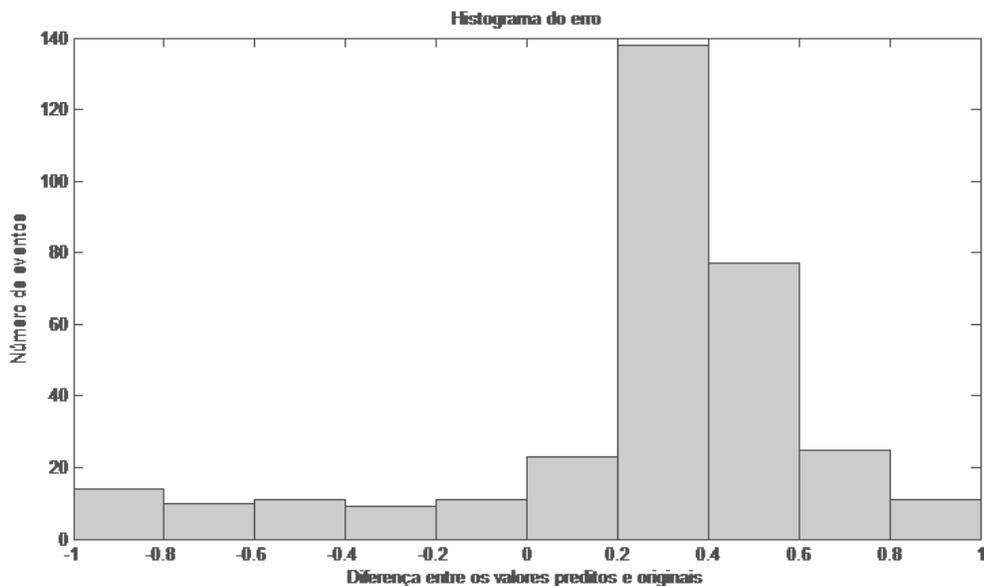


Figura 14 - Cenário 2 - Histograma do erro

4.3.3 Cenário 3: MA com contornos por ECJV

Este cenário trata da aplicação do método MA à série de contorno obtida pelo algoritmo ECJV. A Figura 15 ilustra, dentre dez simulações realizadas, a melhor relação entre a série original e a predita. A Figura 16 apresenta o histograma do erro para a simulação.

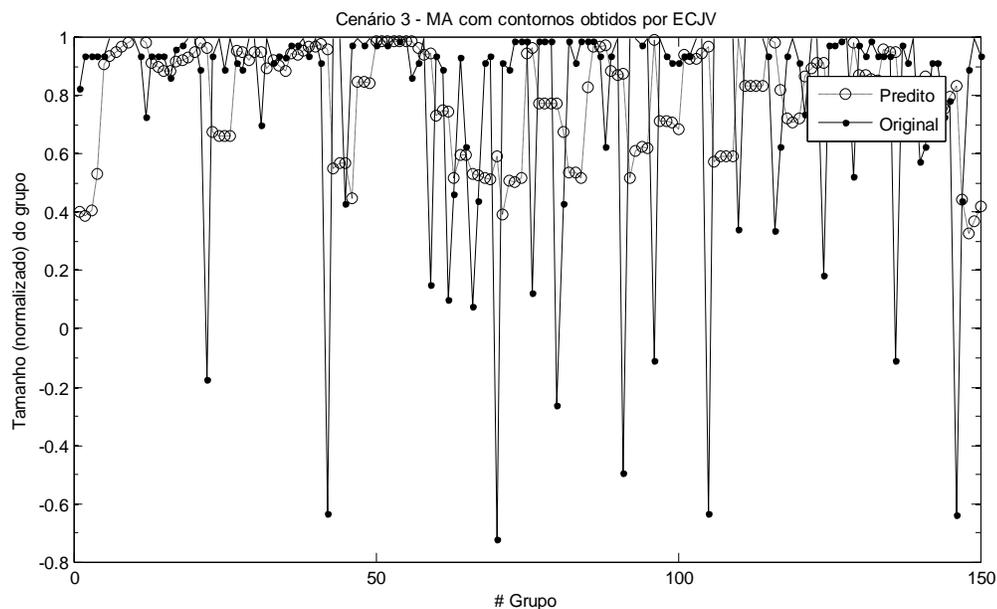


Figura 15 - Cenário 3 - Método MA com os contornos obtidos por ECJV

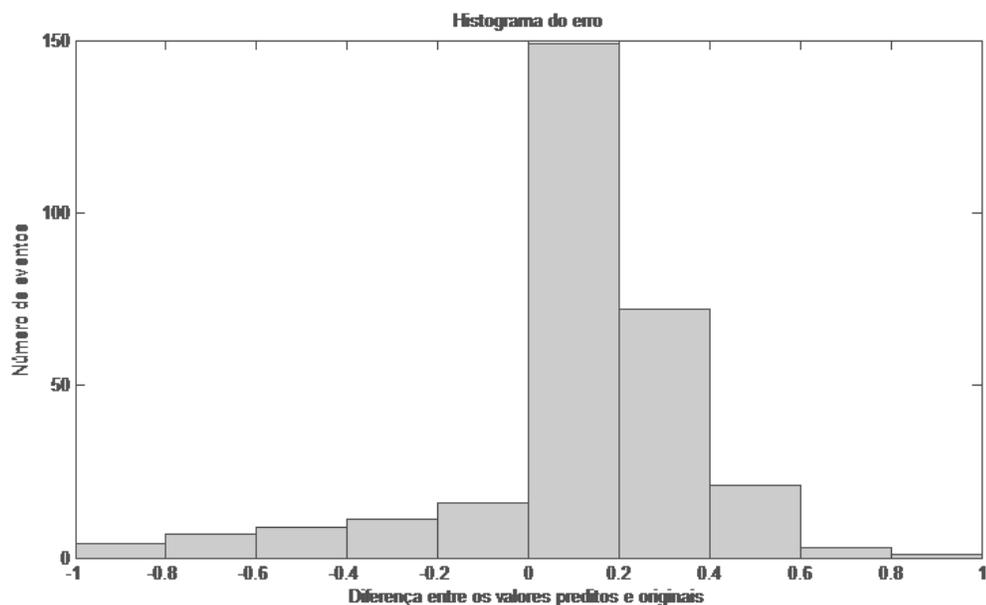


Figura 16 - Cenário 3 - Histograma do erro

Observamos que no Cenário 3 houve uma melhora na predição quando comparado aos Cenários 1 e 2, visto que o histograma de erro está centrado em uma região próxima ao *zero*. Isto indica que a associação do método MA com os contornos obtidos através do algoritmo ECJV apresenta uma boa predição quando comparamos o mesmo método associado aos dados originais e ao contorno por ECJF.

4.3.4 Cenário 4: AES com dados originais

Este cenário trata da aplicação do método AES à série original normalizada. A Figura 17 ilustra, dentre dez simulações realizadas, a melhor relação entre a série original e a predita. A Figura 18 apresenta o histograma do erro. A configuração *weight in smoothing* no Minitab foi ajustado em 0,7.

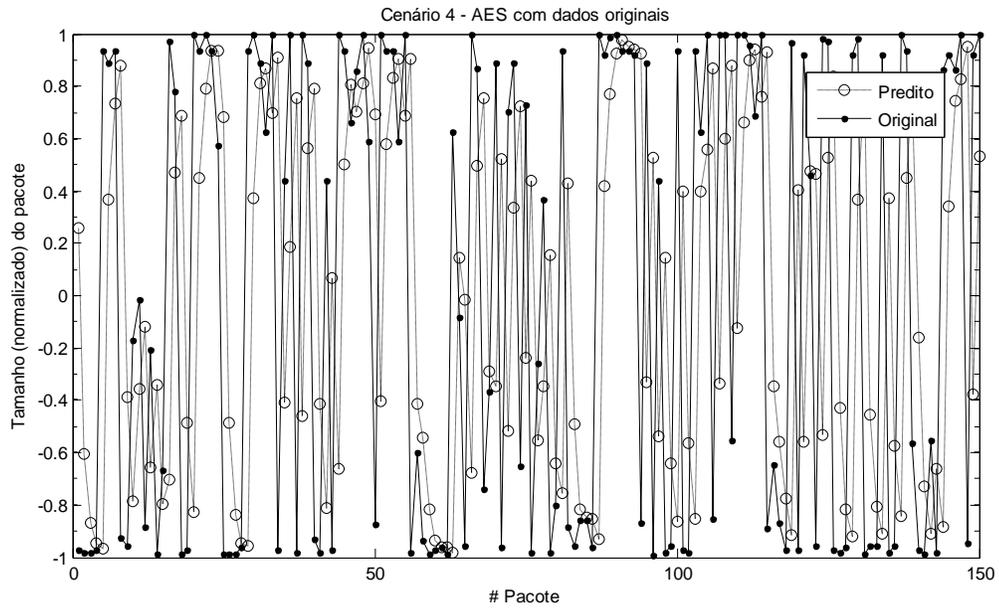


Figura 17 - Cenário 4 - Método AES com os dados originais

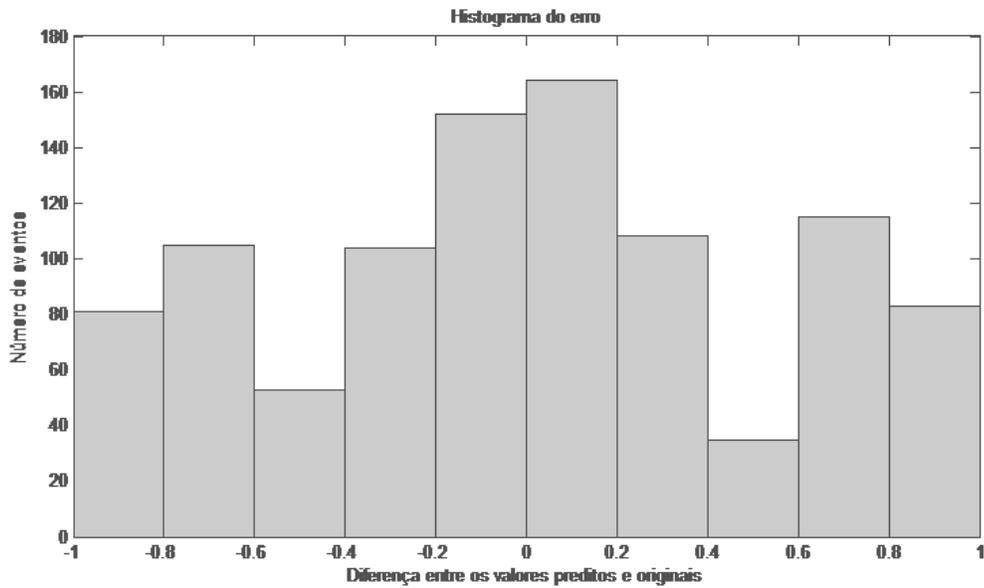


Figura 18 - Cenário 4 - Histograma do erro

No Cenário 4 observa-se que o erro está bastante disperso, indicando que a associação do método AES com os dados originais não é uma abordagem eficiente.

4.3.5 Cenário 5: AES com contornos por ECJF

Este cenário trata da aplicação do método AES à série de contorno obtida pelo algoritmo ECJF. A Figura 19 ilustra a melhor relação entre a série original e a predita, entre 10 simulações realizadas. A Figura 20 apresenta o histograma do erro.

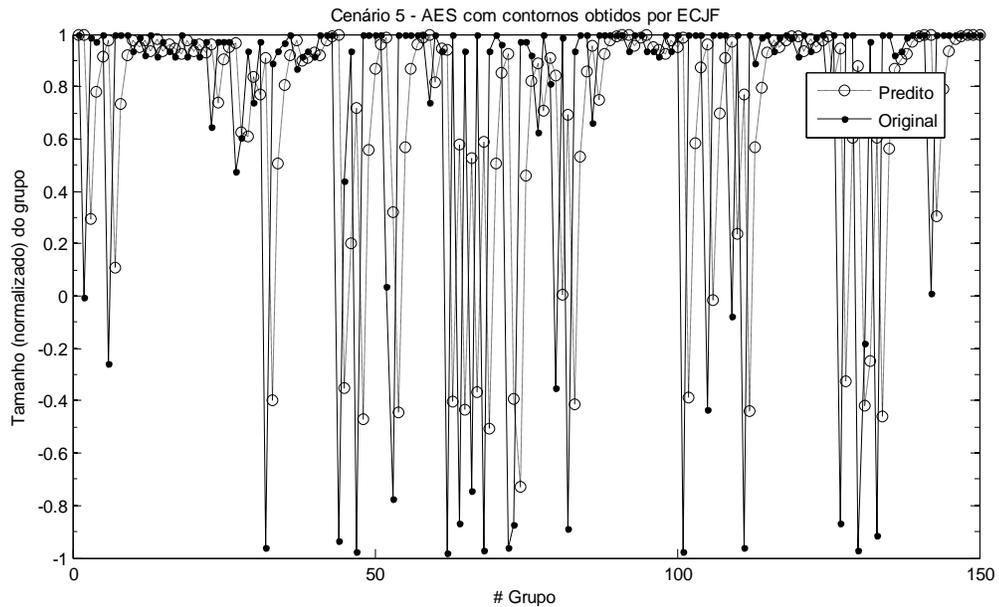


Figura 19 - Cenário 5 - Método AES com contornos obtidos por ECJF

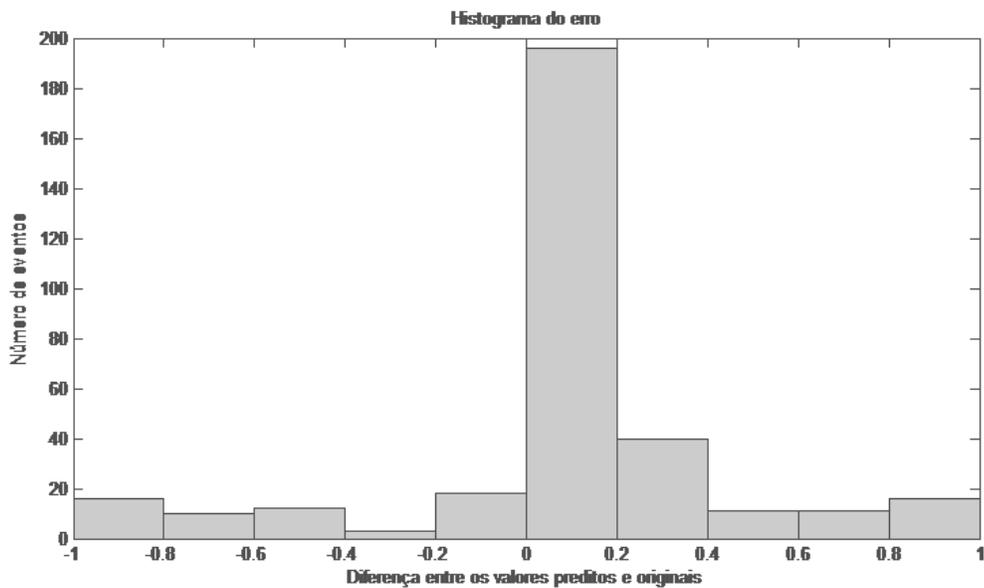


Figura 20 - Cenário 5 - Histograma do erro

Observa-se que no Cenário 5 houve uma melhora na predição, visto que o histograma do erro está concentrado em uma região próxima ao *zero*. Isto indica que a associação do método AES com os contornos obtidos através do algoritmo ECJF apresenta uma boa predição quando comparado ao mesmo método associado aos dados originais.

4.3.6 Cenário 6: AES com contornos por ECJV

Este cenário envolve a aplicação do método AES à série de contorno obtida pelo algoritmo ECJV. A Figura 21 ilustra a melhor relação entre a série original e a predita, dentre 10 simulações realizadas. O histograma do erro é apresentado na Figura 22.

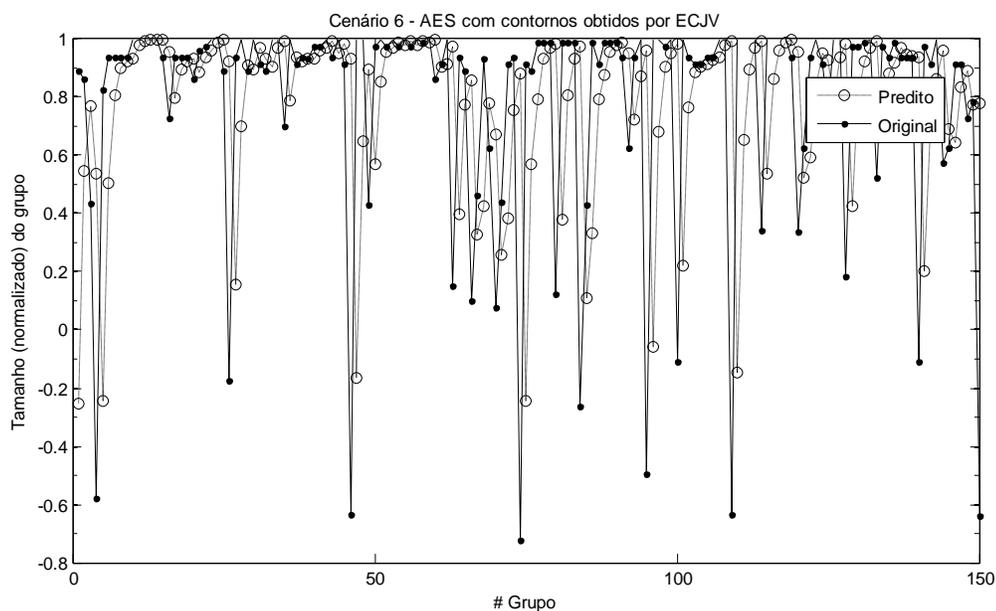


Figura 21 - Cenário 6 - Método AES com contornos obtidos por ECJV

Verifica-se que no Cenário 6 houve uma melhora na predição, visto que o histograma do erro está concentrado em uma região próxima ao 0, sem muitas dispersões, como as presentes no Cenário 5. Isto indica que a associação do método AES com os contornos obtidos através do algoritmo ECJV apresenta uma boa predição quando comparado ao mesmo método associado aos dados originais e aos contornos por ECJF, mostrando que o algoritmo ECJV favorece a uma eficiente predição.

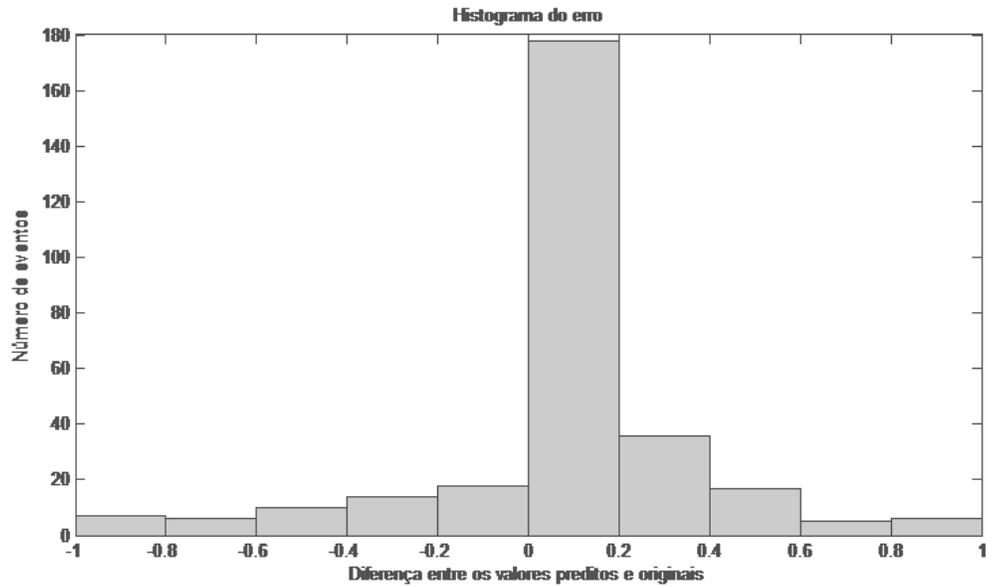


Figura 22 - Cenário 6 - Histograma do erro

4.3.7 Cenário 7: HW com dados originais

Este cenário trata da aplicação do método HW à série original normalizada. A Figura 23 ilustra, dentre dez simulações realizadas, a melhor relação entre a série original e a predita. A Figura 24 apresenta o histograma do erro. A configuração no Minitab foi de $level = 0,5$; $trend = 0,1$ e $seasonal = 0,1$, com $seasonal\ length = 3$.

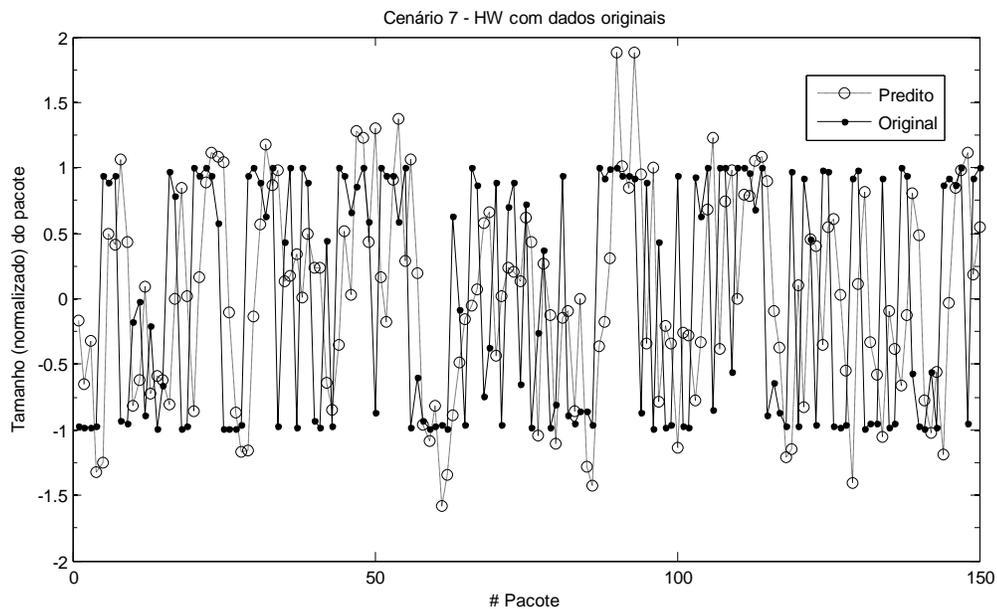


Figura 23 - Cenário 7 - Método HW com os dados originais

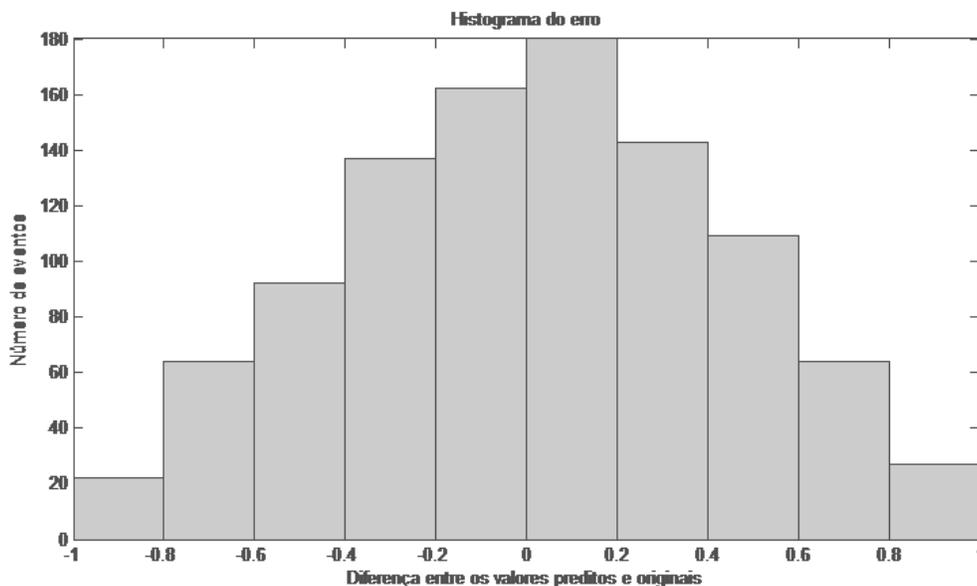


Figura 24 - Cenário 7 - Histograma do erro

No Cenário 7 visualiza-se um histograma de erro bastante distribuído, além de valores preditos bem distintos do valor original. Isto é um indicativo que a associação do método Holt-Winters aos dados originais não é uma abordagem segura.

4.3.8 Cenário 8: HW com contornos por ECJF

Este cenário envolve a aplicação do método HW à série de contorno obtida pelo algoritmo ECJF. A Figura 25 ilustra a melhor relação entre a série original e a predita, entre 10 simulações realizadas. O histograma do erro está apresentado na Figura 26.

No Cenário 8 visualiza-se o histograma do erro um pouco distribuído, embora com uma concentração mais próxima ao zero. Isto indica que a associação do método Holt-Winters aos contornos obtidos pelo algoritmo ECJF não é uma boa prática, mas se mostra superior ao mesmo método quando associado aos dados originais.

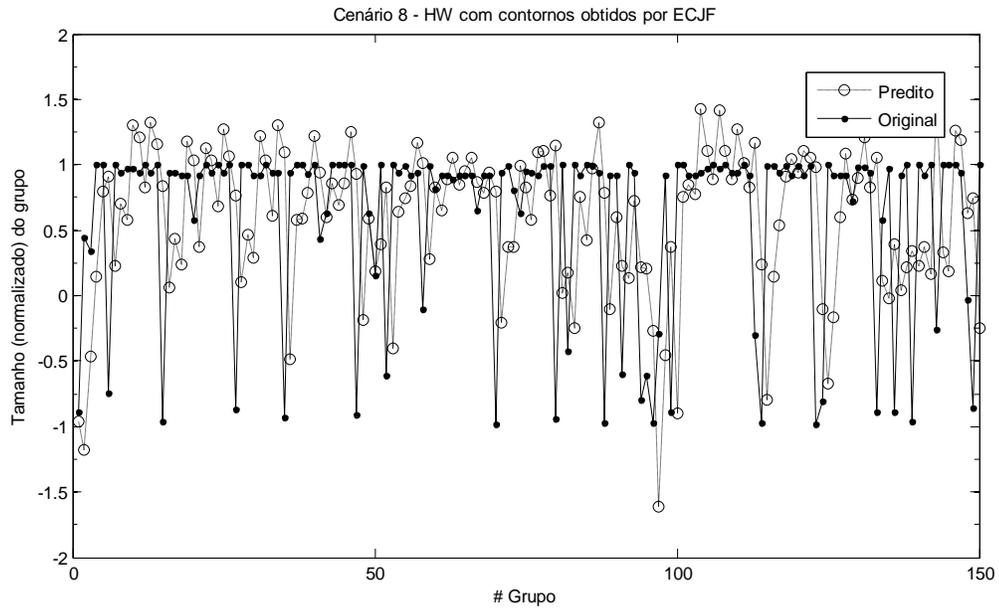


Figura 25 - Cenário 8 - Método HW com contornos obtidos por ECJF

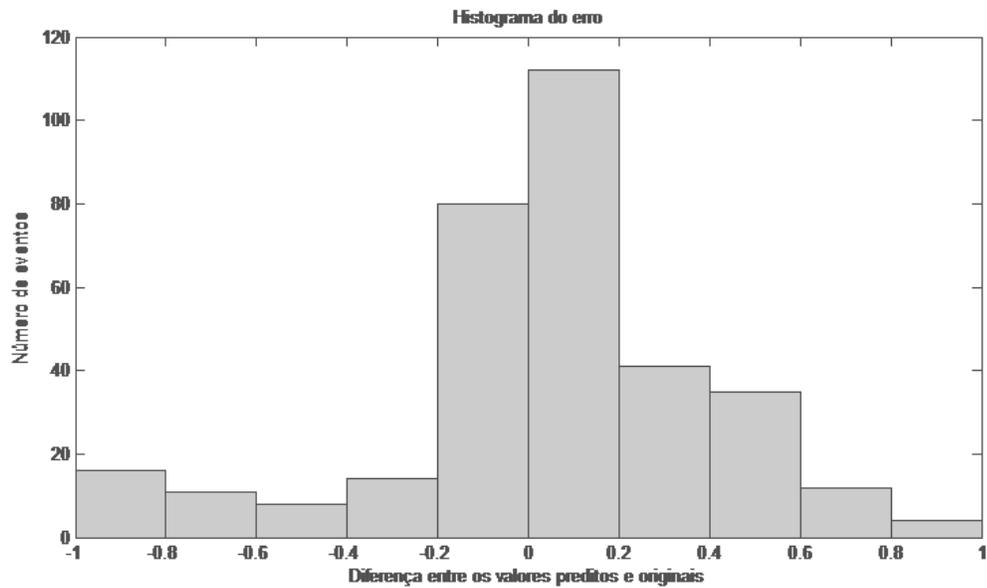


Figura 26 - Cenário 8 - Histograma do erro

4.3.9 Cenário 9: HW com contornos por ECJV

Este cenário trata da aplicação do método HW à série de contorno obtida pelo algoritmo ECJV. A Figura 27 ilustra a melhor relação entre a série original e a predita, dentre 10 simulações realizadas. O histograma do erro está apresentado na Figura 28.

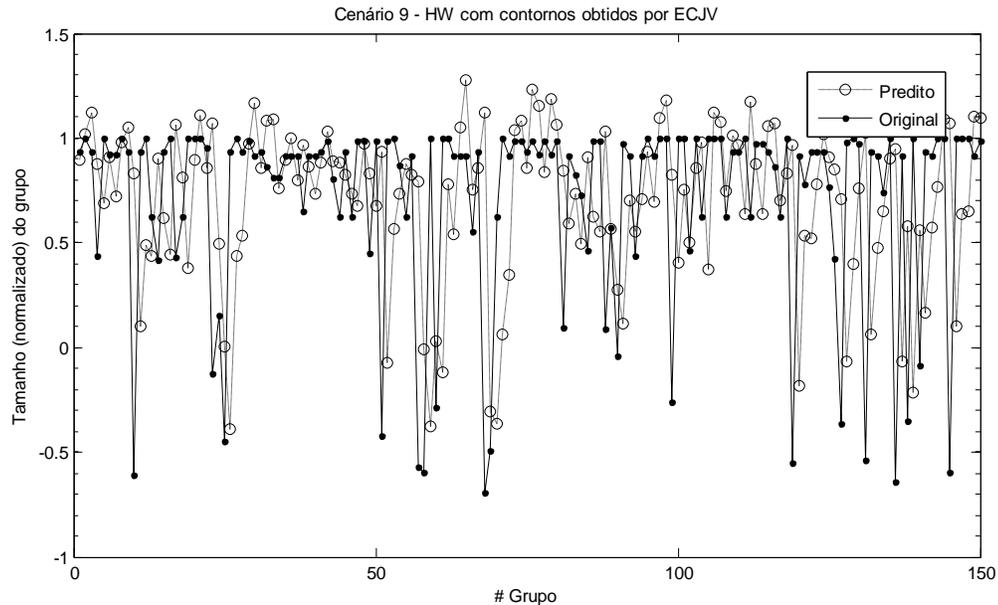


Figura 27 - Cenário 9 - Método HW com contornos obtidos por ECJV

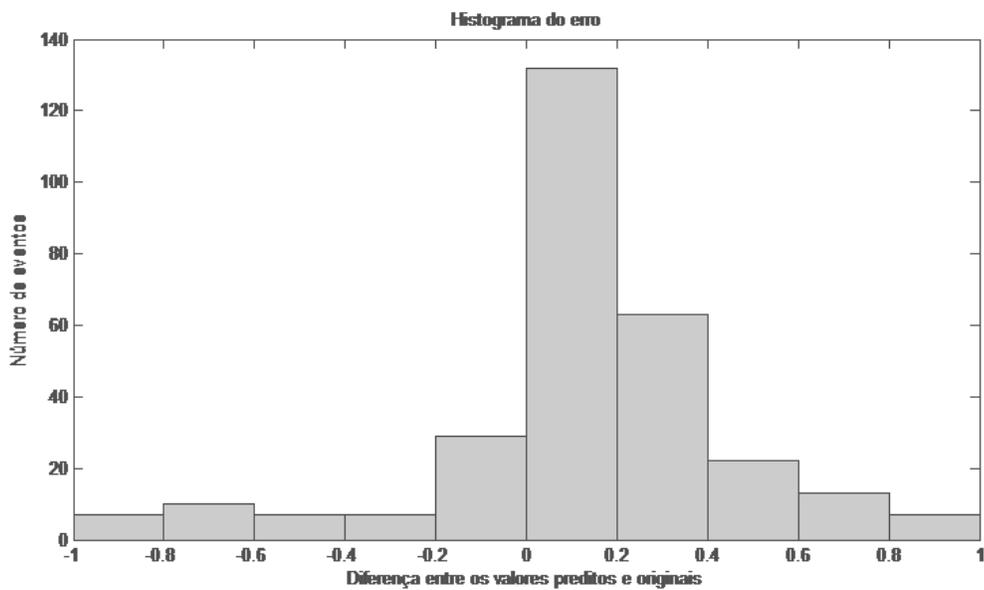


Figura 28 - Cenário 9 - Histograma do erro

De forma similar ao observado no Cenário 8, no Cenário 9 o histograma do erro apresenta-se um pouco distribuído, embora com uma concentração próxima ao *zero*, indicando que a associação do método Holt-Winters aos contornos obtidos pelo algoritmo ECJV não é uma abordagem totalmente eficiente, mostrando-se superior ao mesmo método associado aos dados originais.

4.3.10 Cenário 10: RFC 2988 com dados originais

Este cenário analisa o método RFC 2988 aplicado à série original normalizada. A Figura 29 ilustra, dentre dez simulações realizadas, a melhor relação entre a série original e a predita. O histograma do erro está disponível na Figura 30.

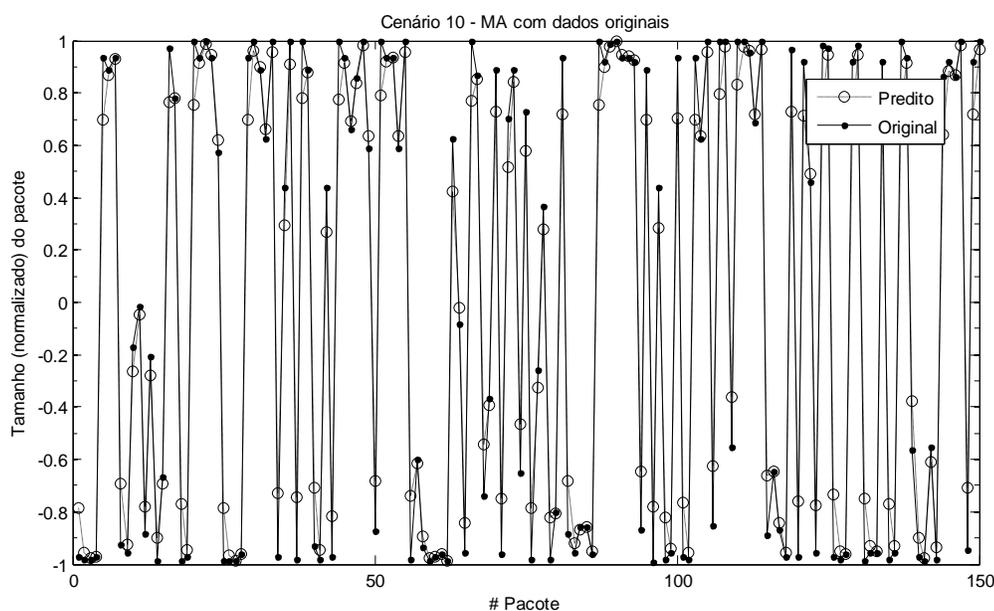


Figura 29 - Cenário 10 - Método RFC 2988 com os dados originais

No Cenário 10 verifica-se um histograma do erro distribuído, indicando que o método estatístico RFC 2988 associado aos dados originais não foi eficiente na predição, contudo, apresenta um melhor desempenho quando comparado aos métodos MA, AES e Holt-Winters.

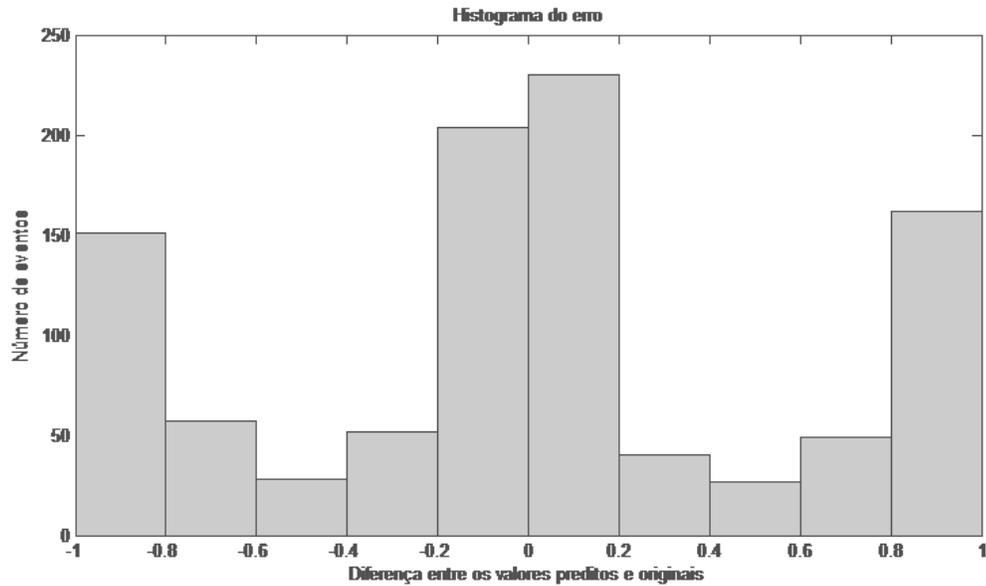


Figura 30 - Cenário 10 - Histograma do erro

4.3.11 Cenário 11: RFC 2988 com contornos por ECJF

Este cenário envolve a aplicação do método RFC 2988 à série de contorno obtida pelo algoritmo ECJF. A Figura 31 ilustra, dentre dez simulações realizadas, a melhor relação entre a série original e a predita. A Figura 32 apresenta o histograma do erro.

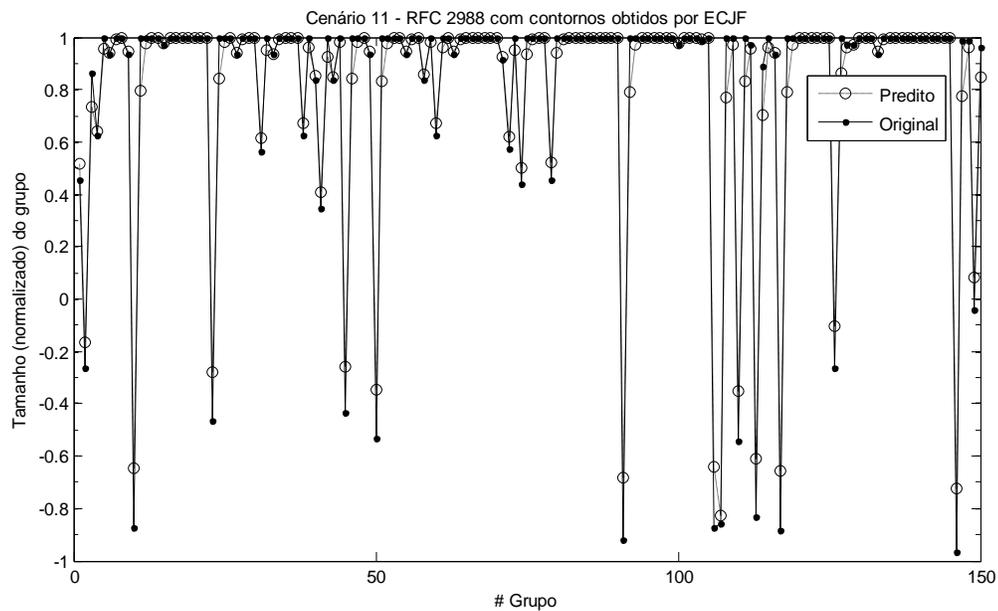


Figura 31 - Cenário 11 - Método RFC 2988 com contornos obtidos por ECJF

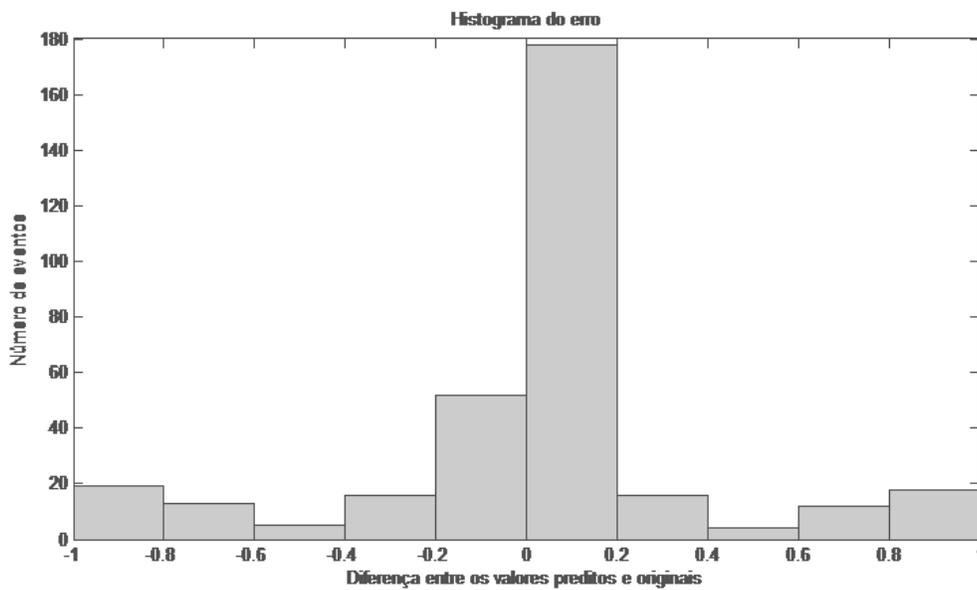


Figura 32 - Cenário 11 - Histograma do erro

Observamos no Cenário 11 uma distribuição do erro bastante concentrada em torno de zero. Isto indica que o método RFC 2988 associado aos contornos obtidos pelo algoritmo ECJF possibilita uma predição eficiente. Quando comparamos com o resultado exposto no Cenário 10, verifica-se que o algoritmo ECJF propicia uma boa atuação do método preditor.

4.3.12 Cenário 12: RFC 2988 com contornos por ECJV

Este cenário trata da aplicação do método RFC 2988 à série de contorno obtida pelo algoritmo ECJV. A Figura 33 ilustra, dentre dez simulações realizadas, a melhor relação entre a série original e a predita. O histograma do erro é apresentado na Figura 34.

Observa-se no Cenário 12 um histograma do erro concentrado em torno de zero. Isto é um indicativo que o método RFC 2988 associado aos contornos obtidos pelo algoritmo ECJV possibilita uma boa predição, mas não melhor que quando utilizado o algoritmo ECJF. Quando se compara com o resultado exposto no Cenário 10, verifica-se que o algoritmo ECJV favorece uma boa atuação do método preditor.

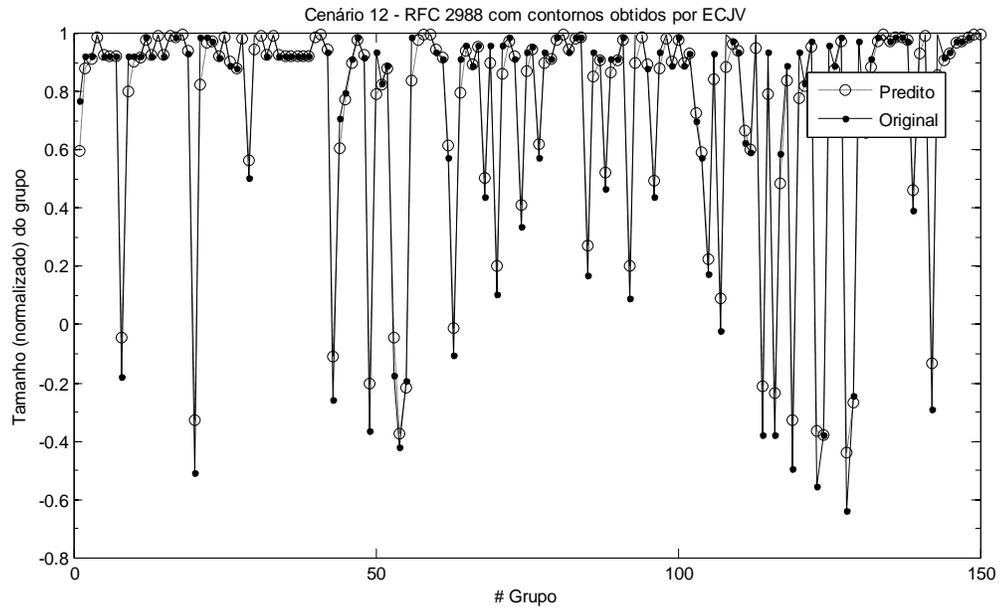


Figura 33 - Cenário 12 - Método RFC 2988 com contornos obtidos por ECJV

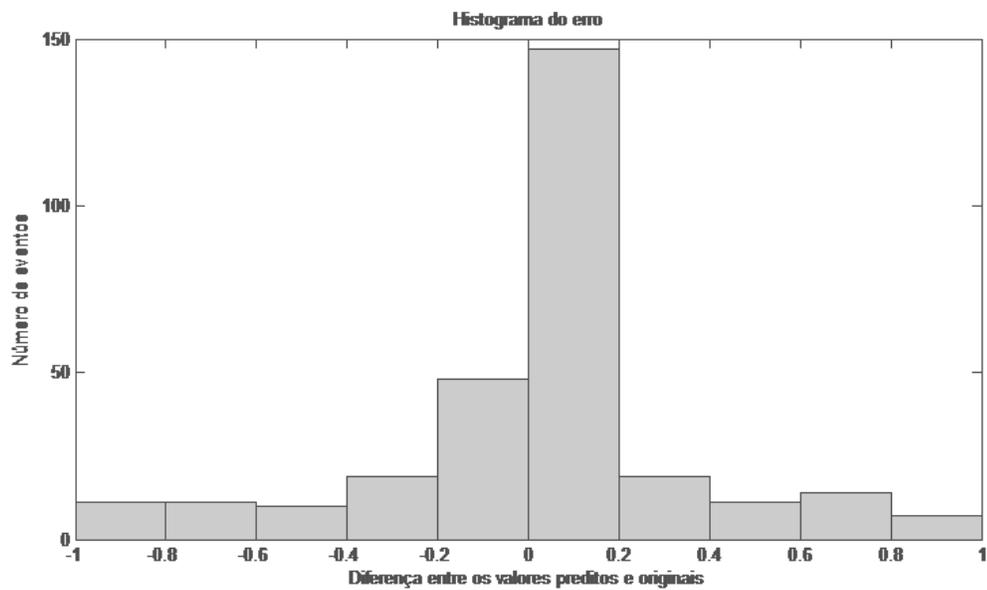


Figura 34 - Cenário 12 - Histograma do erro

4.3.13 Cenário 13: MLP com dados originais

Este cenário estuda o desempenho da rede MLP aplicada à série original normalizada. A Figura 35 ilustra, dentre dez simulações realizadas, a melhor relação entre a série original e a predita. O histograma do erro é apresentado na Figura 36.

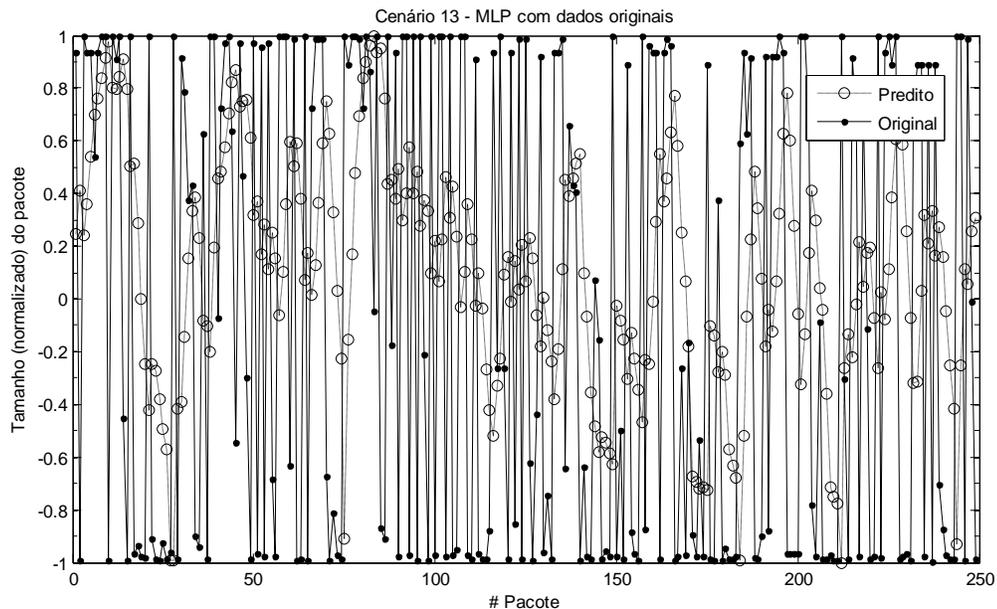


Figura 35 - Cenário 13 - Método MLP com os dados originais

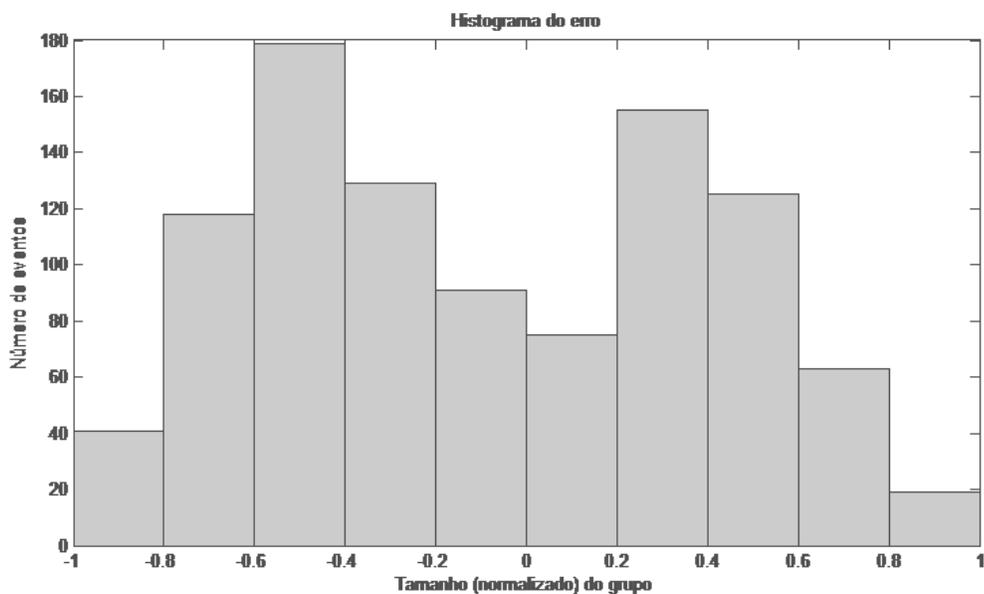


Figura 36 - Cenário 13 - Histograma do erro

No Cenário 13 verifica-se uma distribuição do erro por toda a amplitude dos valores, indicando que a rede MLP associada aos dados originais não resulta em uma abordagem preditiva totalmente eficiente, contudo, apresenta um melhor desempenho quando comparado aos métodos estatísticos MA, AES e Holt-Winters com traços originais.

4.3.14 Cenário 14: MLP com contornos por ECJF

Este cenário envolve a aplicação da rede MLP à série de contorno obtida pelo algoritmo ECJF. A Figura 37 ilustra, dentre dez simulações realizadas, a melhor relação entre a série original e a predita.

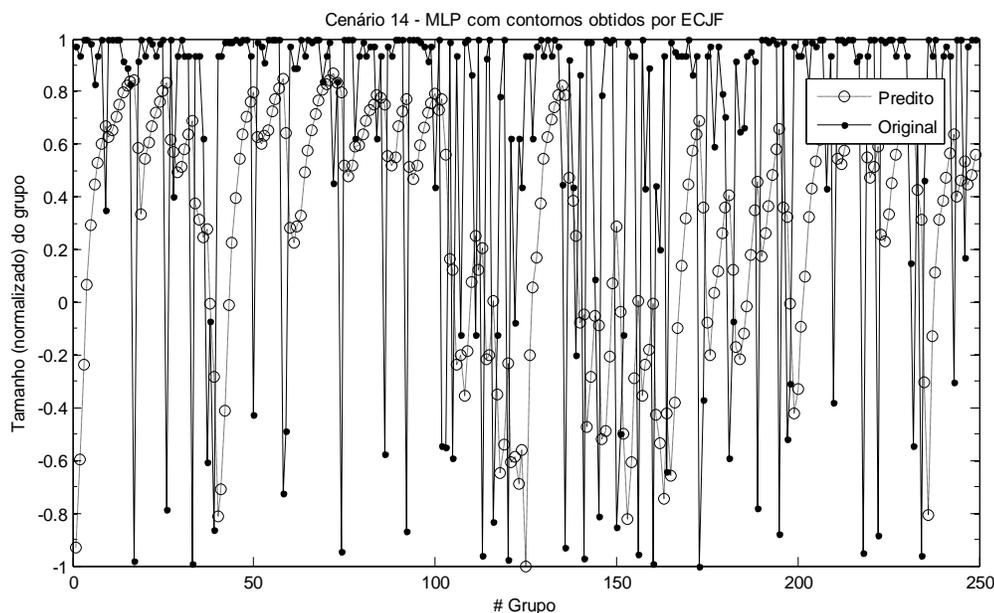


Figura 37 - Cenário 14 - Método MLP com contornos obtidos por ECJF

Através da Figura 38 observa-se o histograma do erro, verificando que o Cenário 14 apresenta o histograma do erro distribuído por toda a amplitude dos valores, mas há presença de alguns pontos concentrados em valores próximos a *zero*, indicando que a rede MLP associada aos contornos obtidos pelo algoritmo ECJF não é totalmente eficiente na predição, mas, apresenta um melhor desempenho quando comparado aos métodos estatísticos MA, AES e Holt-Winters e à própria rede MLP em associação aos dados originais.

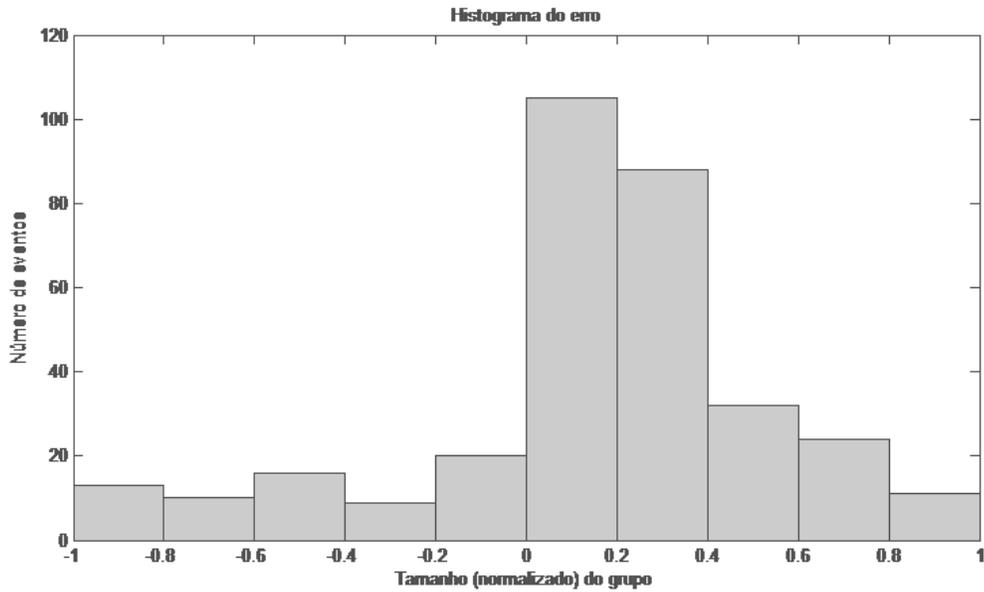


Figura 38 - Cenário 14 - Histograma do erro

4.3.15 Cenário 15: MLP com contornos por ECJV

Este cenário analisa a aplicação da rede MLP à série de contorno obtida pelo algoritmo ECJV. A Figura 39 ilustra, dentre dez simulações realizadas, a melhor relação entre a série original e a predita. A Figura 40 apresenta o histograma do erro.

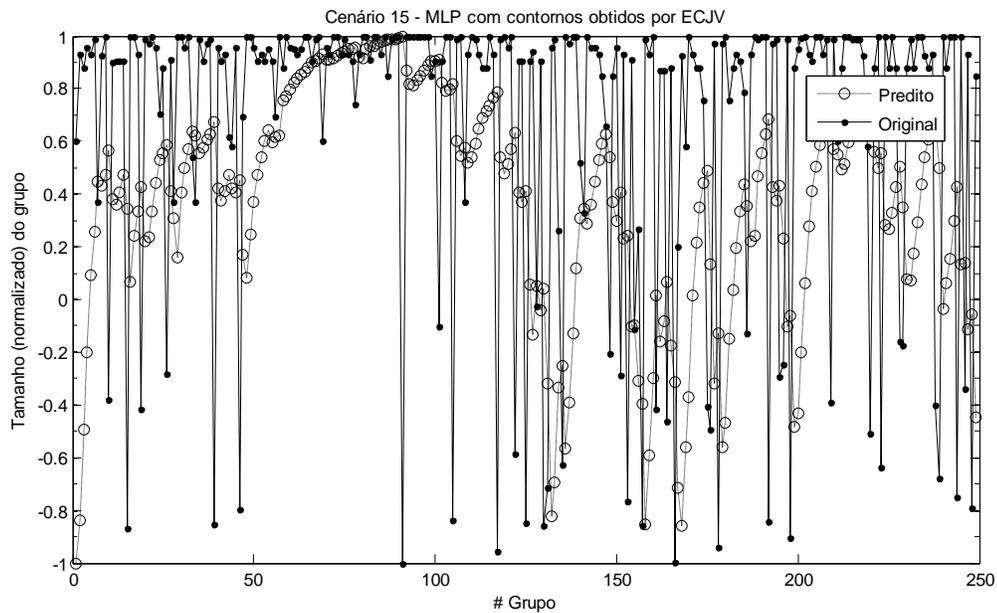


Figura 39 - Cenário 15 - Método MLP com contornos obtidos por ECJV

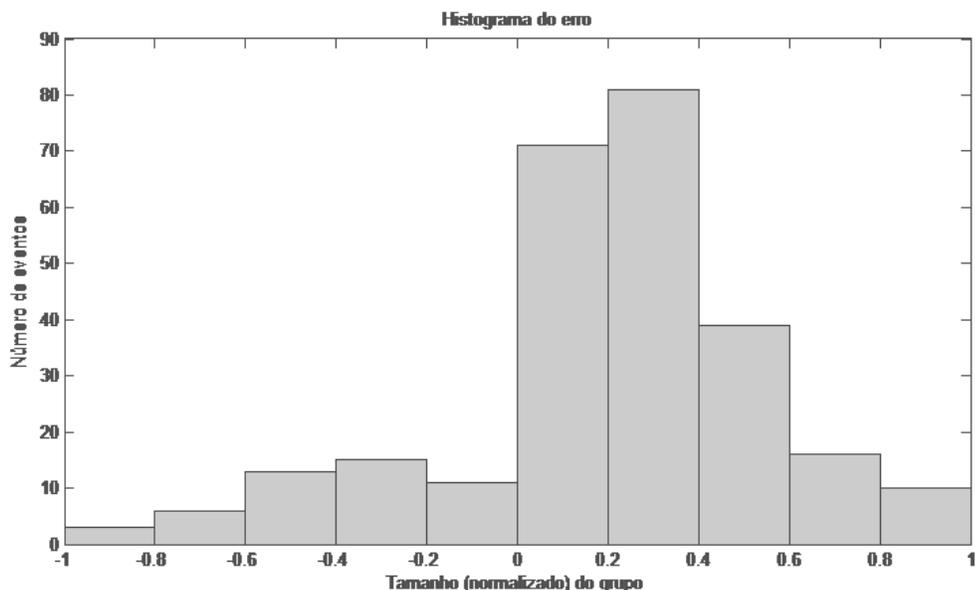


Figura 40 - Cenário 15 - Histograma do erro

Observa-se no Cenário 15 que o erro está concentrado em valores próximos a *zero*, com poucas rajadas de erros em outras faixas de valores. Isto indica que a rede MLP associada aos contornos obtidos pelo algoritmo ECJV é eficiente na predição, apresentando um melhor desempenho quando comparado aos métodos estatísticos MA, AES e Holt-Winters e ao próprio método MLP quando associado aos dados originais e aos contornos por ECJF.

4.3.16 Cenário 16: Narx com dados originais

Este cenário envolve a aplicação da rede Narx à série original normalizada. A Figura 41 ilustra, dentre dez simulações realizadas, a melhor relação entre a série original e a predita. O histograma do erro é apresentado na Figura 42.

No Cenário 16 observa-se o histograma do erro distribuído por toda a amplitude dos valores, com alguns pontos de concentração no valor *zero*, indicando que a rede NARX associada aos dados originais não é totalmente eficiente na predição, contudo, apresenta um melhor desempenho quando comparado aos métodos estatísticos MA, AES e Holt-Winters, e desempenho semelhante ao MLP com os dados originais de entrada.

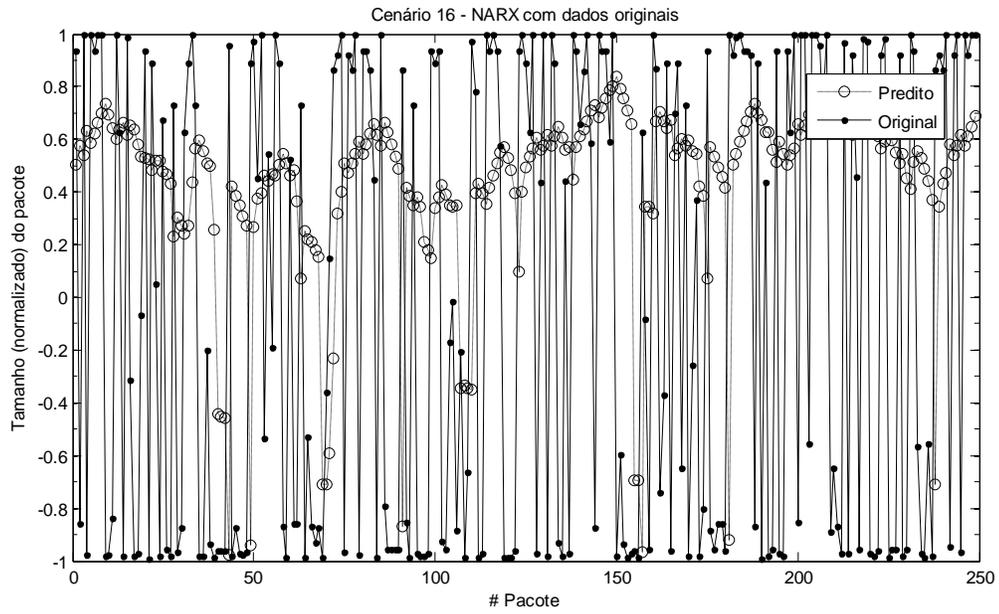


Figura 41 - Cenário 16 - Método Narx com os dados originais

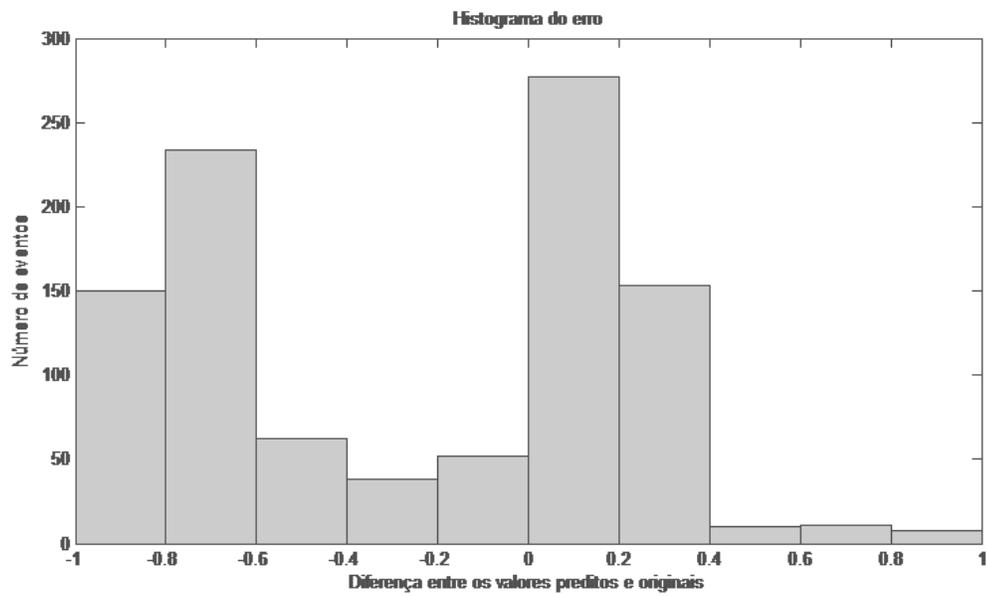


Figura 42 - Cenário 16 - Histograma do erro

4.3.17 Cenário 17: Narx com contornos por ECJF

Este cenário avalia a rede NARX aplicada à série de contorno obtida pelo algoritmo ECJF. A Figura 43 ilustra, dentre dez simulações realizadas, a melhor relação entre a série original e a predita. O histograma do erro está disponível na Figura 44.

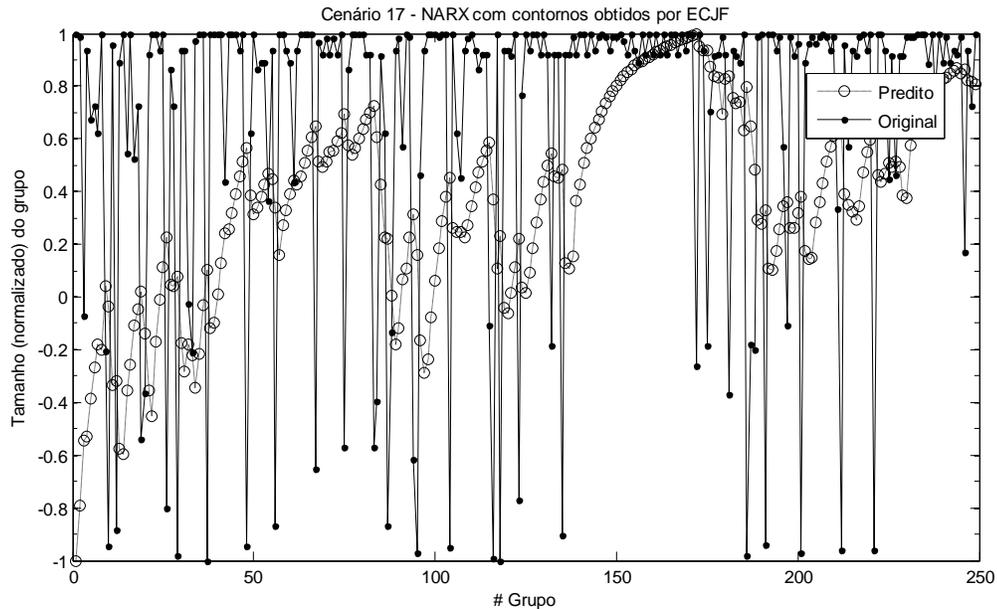


Figura 43 - Cenário 17 - Método Narx com contornos obtidos por ECJF

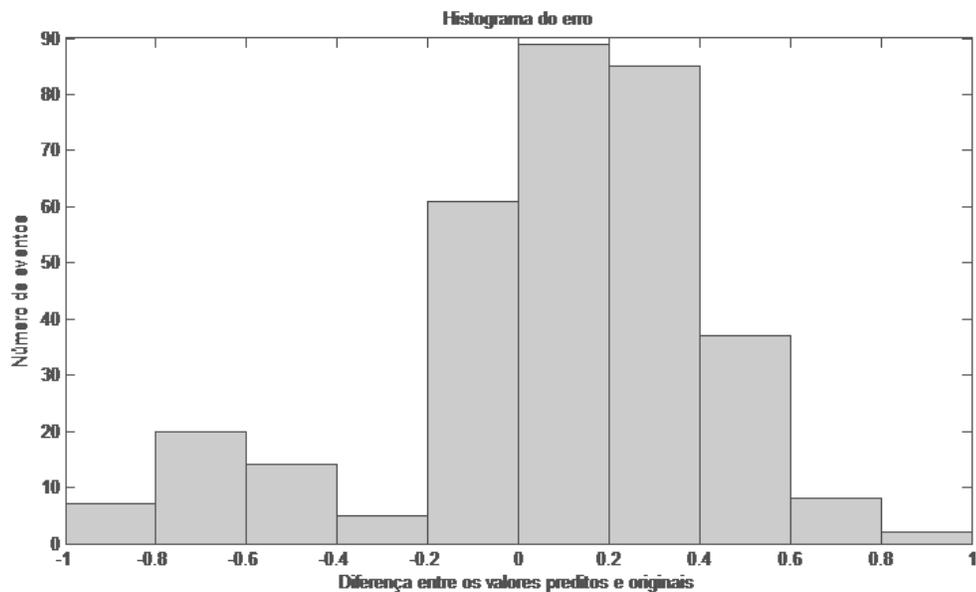


Figura 44 - Cenário 17 - Histograma do erro

Observa-se no Cenário 17 uma distribuição do erro concentrada em valores próximos a *zero*, com algumas rajadas de erro. Isto é um indicativo de que a rede NARX associada aos contornos obtidos pelo algoritmo ECJF é eficiente na predição, apresentando um melhor desempenho quando comparado aos métodos estatísticos MA, AES e Holt-Winters, ao MLP e ao próprio método NARX quando associado aos dados originais. Coaduna-se, portanto, a eficiência dos algoritmos extratores de contorno no auxílio a uma predição eficiente.

4.3.18 Cenário 18: Narx com contornos por ECJV

Este cenário analisa o desempenho da rede Narx aplicada à série de contorno obtida pelo algoritmo ECJV. A Figura 45 apresenta, dentre dez simulações realizadas, a melhor relação entre a série original e a predita. O histograma do erro pode ser visto na Figura 46.

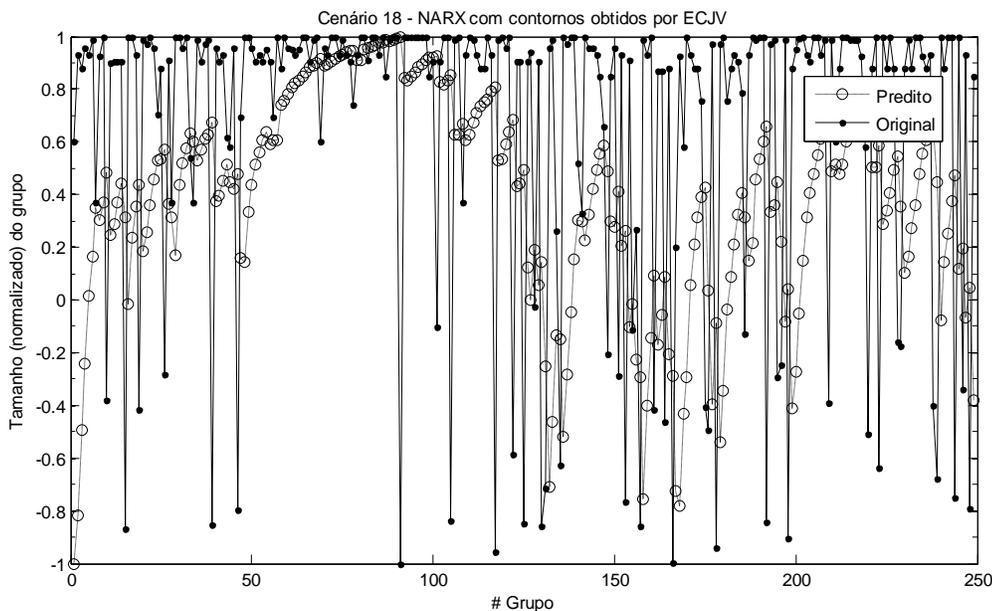


Figura 45 - Cenário 18 - Método Narx com contornos obtidos por ECJV

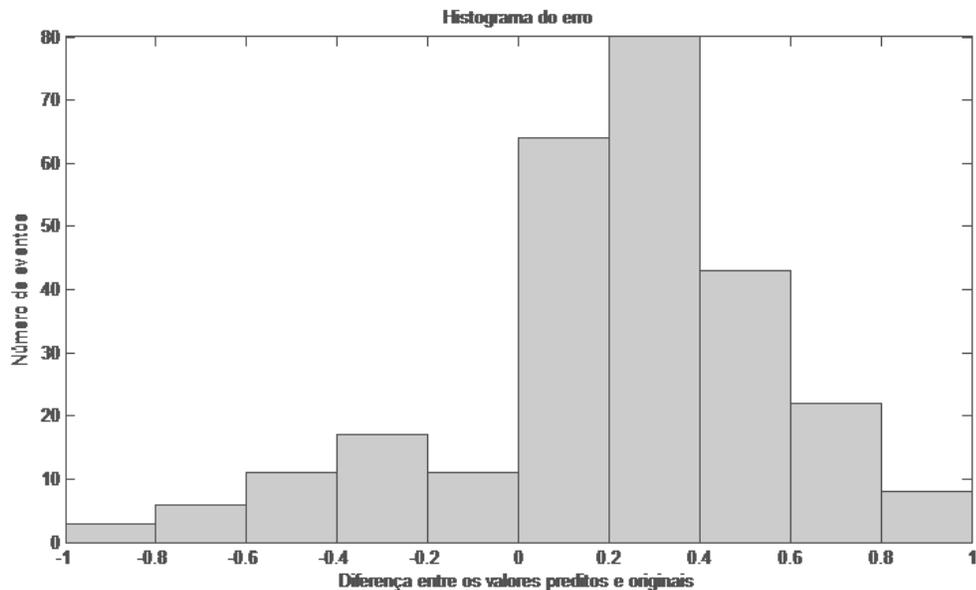


Figura 46 - Cenário 18 - Histograma do erro

O Cenário 18, de forma semelhante ao observado no Cenário 17 apresenta uma distribuição do erro concentrada em valores próximos a *zero*, com algumas rajadas de erro. Isto indica que o método conexionista NARX associado aos contornos obtidos pelo algoritmo ECJV é eficiente na predição, apresentando um melhor desempenho quando comparado aos métodos estatísticos MA, AES e Holt-Winters, à rede MLP e à própria rede NARX quando associada aos dados originais. Conclui-se, portanto, que a associação de modelos neurais com os algoritmos extratores de contorno resultam em uma predição eficiente.

4.3.19 Conjunto de simulações com os dados originais

Conforme exposto na Seção 4.1, foram realizadas dez simulações para cada um dos cenários, resultando em 60 simulações para os cenários que utilizam os dados originais. Nas Tabelas 13 e 14 são representados, respectivamente, o MSE e MAC dos métodos em cada uma das simulações. As simulações estão sinalizadas por $S1, S2, S3, \dots, S9$ e $S10$.

Tabela 13 - MSE das simulações utilizando-se dados originais

MSE	MA	AES	HW	RFC 2988	MLP	NARX
S1	0,9575	1,1709	1,2103	0,0211	0,1999	0,1975
S2	0,9691	1,1811	1,2563	0,0214	0,2036	0,2027
S3	0,9613	1,2071	1,2642	0,0220	0,2031	0,2023
S4	0,9783	1,2077	1,2624	0,0218	0,2069	0,2063
S5	0,9871	1,2315	1,2928	0,0224	0,2057	0,2069
S6	0,9702	1,2143	1,3011	0,0219	0,2062	0,2063
S7	0,9904	1,2265	1,3123	0,0221	0,2061	0,2079
S8	0,9922	1,2730	1,3290	0,0231	0,2082	0,2099
S9	0,9523	1,1720	1,2414	0,0213	0,2084	0,2052
S10	0,9256	1,1381	1,2070	0,0205	0,2012	0,2035
MEDIA	0,9684	1,2022	1,2677	0,0218	0,2049	0,2048

Tabela 14 - MAC das simulações utilizando-se dados originais

MAC	MA	AES	HW	RFC 2988	MLP	NARX
S1	0,7421	0,3369	0,3275	0,1513	-0,1078	-0,0078
S2	0,7505	0,3483	0,3529	0,1566	-0,0806	-0,0020
S3	0,7551	0,3501	0,3550	0,1578	-0,0752	-0,0130
S4	0,7523	0,3349	0,3193	0,1494	-0,0024	-0,0066
S5	0,7629	0,3423	0,3425	0,1527	-0,0322	-0,0040
S6	0,7551	0,3344	0,3463	0,1494	0,0199	-0,0078
S7	0,7604	0,3414	0,3470	0,1542	-0,0575	-0,0141
S8	0,7648	0,3410	0,3539	0,1516	-0,0031	-0,0033
S9	0,7515	0,3421	0,3296	0,1537	0,0111	-0,0083
S10	0,7419	0,3333	0,3224	0,1472	0,0243	0,0008
MEDIA	0,7537	0,3405	0,3396	0,1524	-0,0304	-0,0066

Baseando-se pela Tabela 13 apresentada, na Figura 47 é representado o valor médio do MSE dos valores preditos com Intervalo de Confiança de 95%; de forma análoga, baseando-se pela Tabela 14 apresentada, o valor médio do MAC dos valores preditos com Intervalo de Confiança⁹ de 95% é demonstrado na Figura 48.

⁹ Utilizando-se a distribuição t-student e conceitos de grau de liberdade

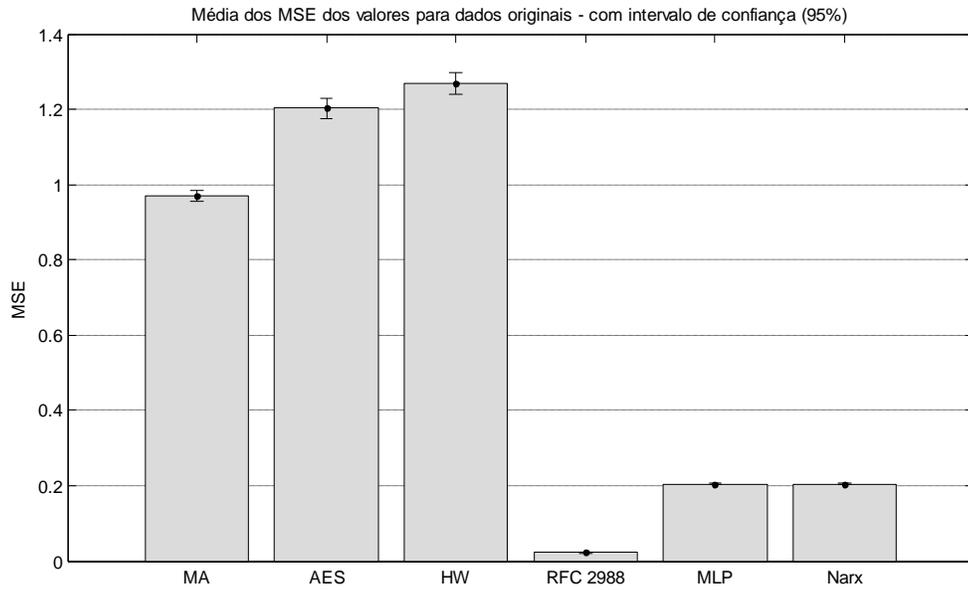


Figura 47 - Média do MSE dos valores preditos para dados originais

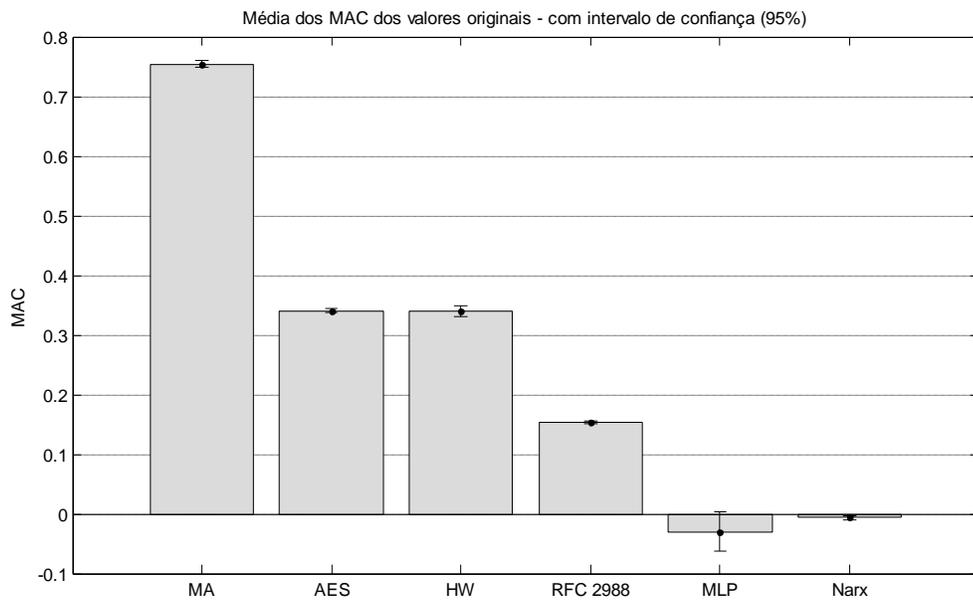


Figura 48 - Média dos MAC dos valores preditos para dados originais

4.3.20 Conjunto de simulações com os contornos (ECJF)

Foram realizadas dez simulações para cada um dos cenários, o que resulta em 60 simulações para os cenários que utilizam os contornos extraídos pelo algoritmo ECJF. Nas Tabelas 15 e 16 são representados, respectivamente, o MSE e MAC dos métodos em cada uma das simulações ocorridas. As simulações estão sinalizadas por $S_1, S_2, S_3, \dots, S_9$ e S_{10} .

Tabela 15 - MSE das simulações utilizando-se contornos por ECJF

MSE	MA	AES	HW	RFC 2988	MLP	NARX
S1	0,5885	0,6776	0,7263	0,0121	0,1274	0,1268
S2	0,5562	0,6592	0,7064	0,0120	0,1183	0,1176
S3	0,4798	0,5666	0,6045	0,0102	0,1030	0,1025
S4	0,4414	0,5422	0,5572	0,0100	0,0966	0,0966
S5	0,4419	0,5197	0,5341	0,0093	0,0980	0,0980
S6	0,4067	0,4981	0,5285	0,0089	0,0906	0,0903
S7	0,5106	0,6287	0,6507	0,0114	0,1078	0,1070
S8	0,3468	0,4655	0,4550	0,0086	0,0800	0,0798
S9	0,3924	0,5005	0,5252	0,0092	0,0892	0,0890
S10	0,3995	0,4745	0,5125	0,0085	0,0848	0,0843
MEDIA	0,4564	0,5533	0,5800	0,0100	0,0996	0,0992

Tabela 16 - MAC das simulações utilizando-se contornos por ECJF

MAC	MA	AES	HW	RFC 2988	MLP	NARX
S1	0,7194	0,3373	0,2454	0,1524	0,0061	0,0029
S2	0,7383	0,3433	0,2346	0,1552	0,0108	0,0080
S3	0,7406	0,3514	0,2120	0,1588	0,0070	0,0044
S4	0,7500	0,3519	0,1754	0,1594	0,0107	0,0092
S5	0,7232	0,3290	0,0770	0,1494	0,0126	0,0112
S6	0,7549	0,3320	0,1931	0,1443	0,0050	0,0036
S7	0,7560	0,3444	0,2074	0,1549	0,0047	0,0028
S8	0,7897	0,3552	0,1755	0,1580	0,0139	0,0132
S9	0,7675	0,3444	0,2155	0,1554	0,0140	0,0128
S10	0,7256	0,3378	0,2217	0,1534	0,0143	0,0131
MEDIA	0,7675	0,3427	0,1958	0,1541	0,0099	0,0081

Baseando-se pela Tabela 15 previamente apresentada, na Figura 49 é representado o valor médio do MSE dos valores preditos com Intervalo de Confiança de 95%, e na Figura 50, é representado o valor médio do MAC dos valores preditos com Intervalo de Confiança de 95% a partir da Tabela 16.

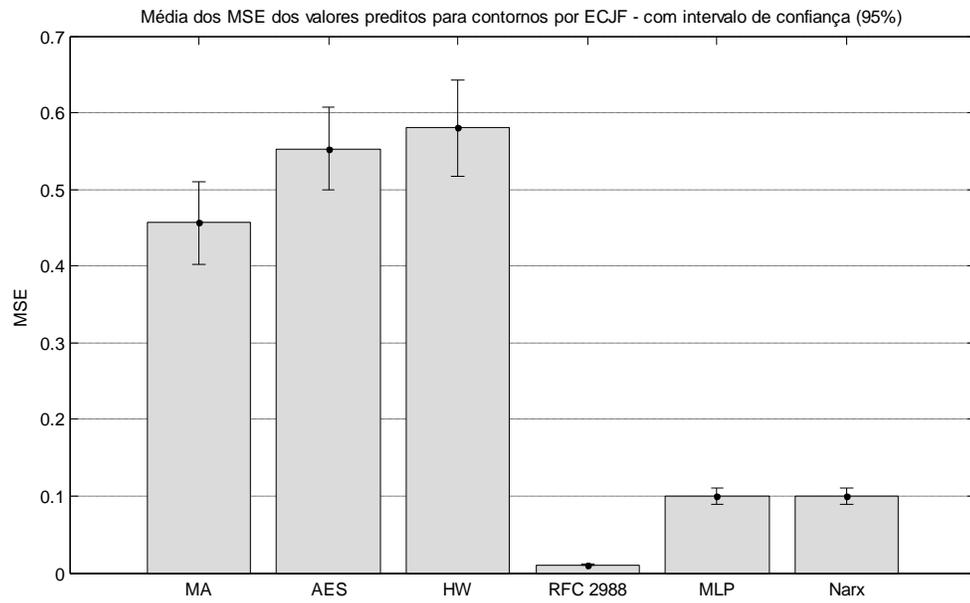


Figura 49 - Média do MSE dos valores preditos para contorno por ECJF

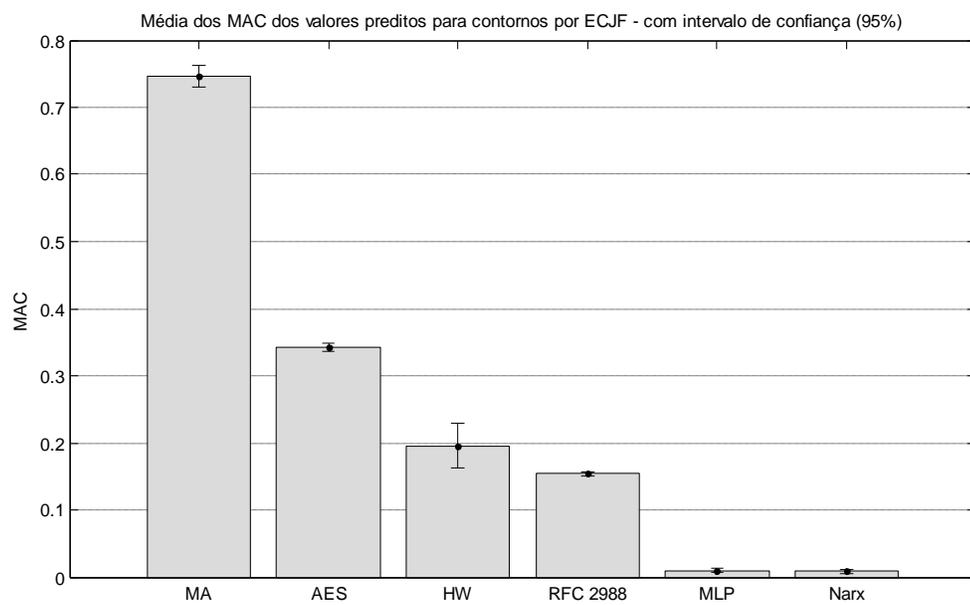


Figura 50 - Média do MAC dos valores preditos para contorno por ECJF

4.3.21 Conjunto de simulações com os contornos (ECJV)

Conforme apresentado na Seção 4.1, foram realizadas dez simulações para cada um dos cenários, o que resulta em 60 simulações para aqueles que utilizam os contornos extraídos pelo algoritmo ECJV. Nas Tabelas 17 e 18 são representados, respectivamente, o MSE e MAC dos métodos em cada uma das simulações, estas sinalizadas por S1, S2, S3, ..., S9 e S10.

Tabela 17 - MSE das simulações utilizando-se contornos por ECJV

MSE	MA	AES	HW	RFC 2988	MLP	NARX
S1	0,3470	0,4179	0,4454	0,0075	0,1087	0,1073
S2	0,4250	0,5080	0,5411	0,0091	0,1169	0,1152
S3	0,3468	0,4292	0,4631	0,0078	0,0982	0,0970
S4	0,1785	0,2291	0,2419	0,0042	0,0594	0,0591
S5	0,2274	0,2761	0,2852	0,0050	0,0570	0,0567
S6	0,2524	0,3086	0,3171	0,0056	0,0625	0,0622
S7	0,3335	0,4333	0,4592	0,0079	0,0786	0,0779
S8	0,2378	0,2805	0,2950	0,0051	0,0682	0,0677
S9	0,2879	0,3636	0,3755	0,0066	0,0884	0,0878
S10	0,2476	0,2811	0,3083	0,0050	0,0753	0,0744
MEDIA	0,2884	0,3527	0,3732	0,0064	0,0813	0,0805

Tabela 18 - MAC das simulações utilizando-se contornos por ECJV

MAC	MA	AES	HW	RFC 2988	MLP	NARX
S1	0,7460	0,3566	0,2980	0,1622	0,0043	0,0007
S2	0,7505	0,3404	0,3187	0,1532	0,0061	0,0017
S3	0,7434	0,3470	0,2653	0,1594	0,0088	0,0060
S4	0,7449	0,3393	0,2525	0,1538	0,0100	0,0090
S5	0,7455	0,3535	0,2283	0,1564	0,0060	0,0049
S6	0,7430	0,3590	0,2463	0,1565	0,0046	0,0035
S7	0,7598	0,3532	0,2937	0,1539	0,0006	-0,0015
S8	0,7343	0,3403	0,2104	0,1504	0,0075	0,0059
S9	0,7449	0,3415	0,2544	0,1512	0,0070	0,0044
S10	0,7238	0,3259	0,2117	0,1481	0,0123	0,0107
MEDIA	0,7436	0,3457	0,2579	0,1545	0,0067	0,0045

Baseando-se pela Tabela 17 previamente apresentada, na Figura 51 é representado o valor médio do MSE dos valores preditos com Intervalo de Confiança de 95%, e na Figura 52, é representado o valor médio do MAC dos valores preditos com Intervalo de Confiança de 95% a partir da Tabela 18.

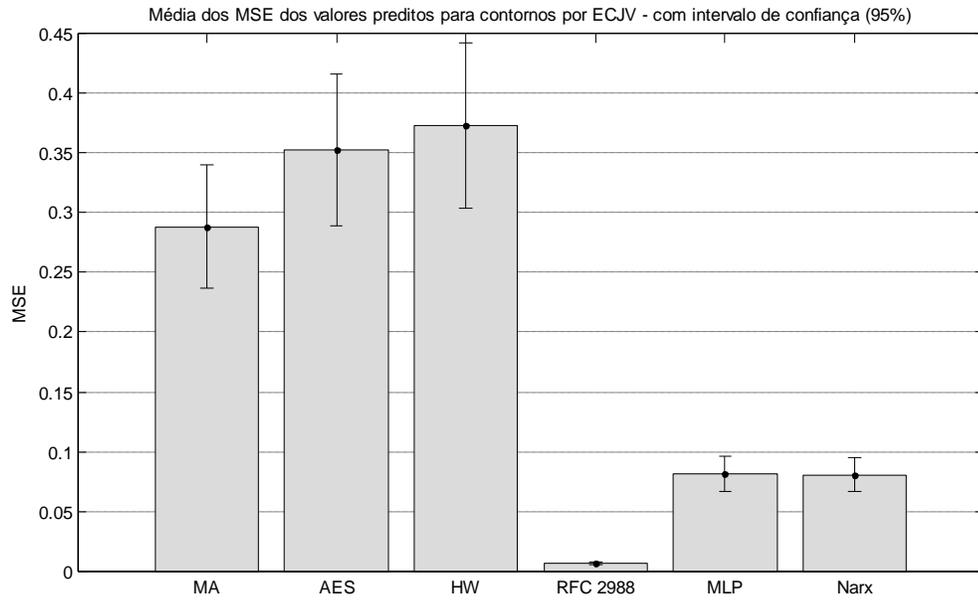


Figura 51 - Média do MSE dos valores preditos para contorno por ECJV

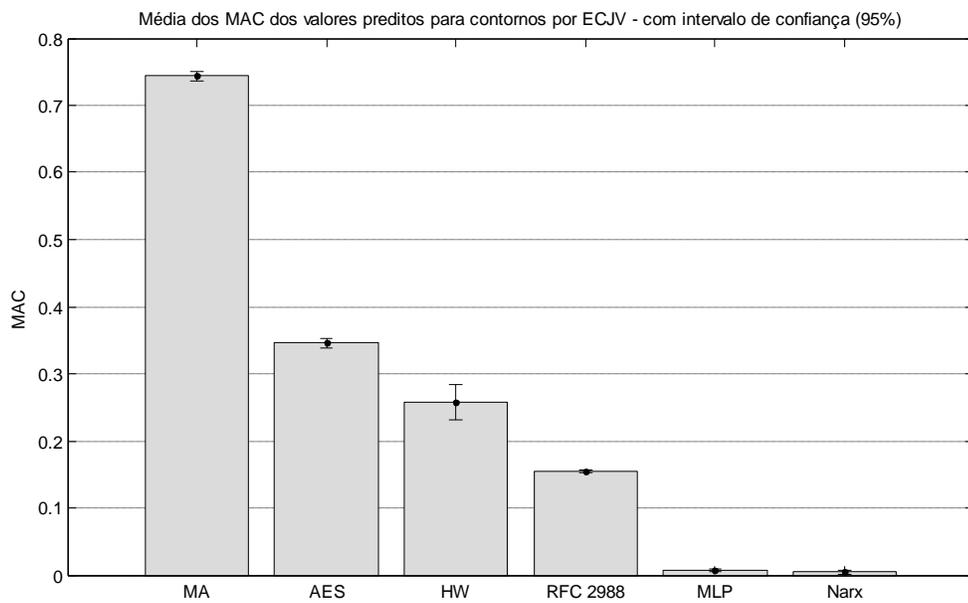


Figura 52 - Média do MAC dos valores preditos para contorno por ECJV

4.4 Análise dos resultados

Fazendo uso do parâmetro estatístico MSE e da Métrica das Amplitudes Conjugadas - MAC, proposta neste trabalho, é realizado o confronto entre os resultados obtidos através da predição estatística e neural.

Na Tabela 19 são apresentados os valores de MSE médios obtidos para os métodos MA, AES, Holt-Winters, RFC 2988, MLP e NARX, para os contornos obtidos com uso do algoritmo ECJF e ECJV e os originais normalizados.

Tabela 19 - MSE médio dos métodos preditores

	MA	AES	HW	RFC 2988	MLP	NARX	Média
Originais	0,9684	1,2022	1,2677	0,0218	0,3970	0,2049	0,6770
ECJF	0,4564	0,5533	0,5800	0,0100	0,0996	0,0992	0,2998
ECJV	0,2884	0,3527	0,3732	0,0064	0,0813	0,0805	0,1971

Podemos observar o referido resultado através do gráfico ilustrado na Figura 53:

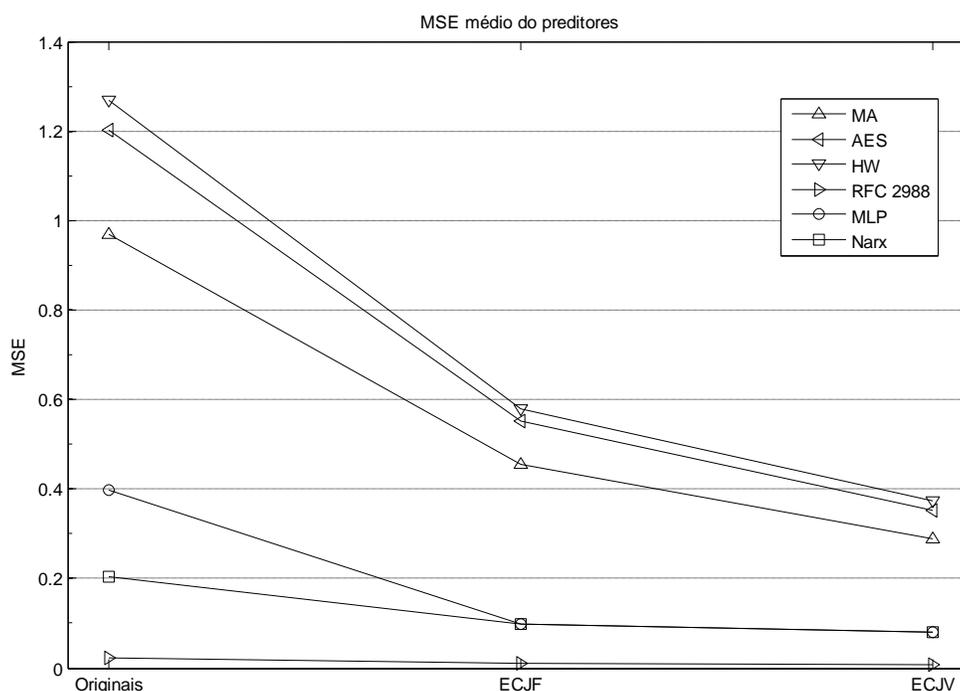


Figura 53 - MSE médio dos métodos preditores

Na Tabela 20 são apresentados os valores de MAC obtidos para os métodos MA, AES, Holt-Winters, RFC 2988, MLP e NARX:

Tabela 20 - MAC médio dos métodos preditores

	MA	AES	HW	RFC 2988	MLP	NARX	Média
Originais	0,7537	0,3405	0,3396	0,1524	-0,0304	-0,0066	0,2582
ECJF	0,7675	0,3427	0,1958	0,1541	0,0099	0,0081	0,2463
ECJV	0,7436	0,3457	0,2579	0,1545	0,0067	0,0045	0,2522

Podemos observar o referido resultado através do gráfico ilustrado na Figura 54:

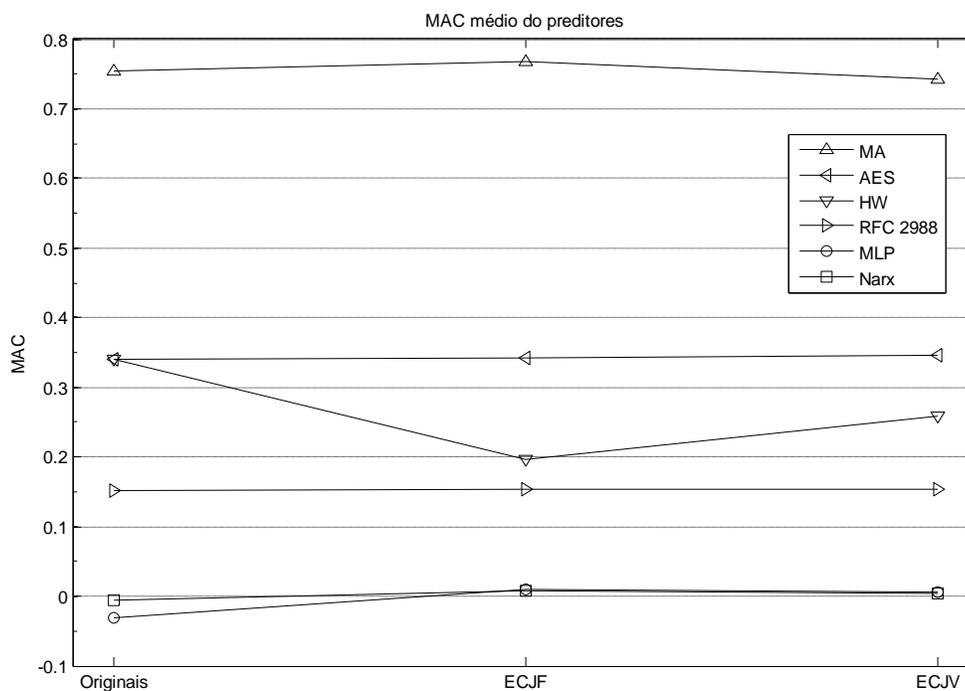


Figura 54 - MAC médio dos métodos preditores

A partir dos dados dispostos na Seção 4.3, observa-se que considerando isoladamente o parâmetro MSE o método RFC 2988 apresentou melhor desempenho, seguido de Narx, MLP, MA, AES e Holt-Winters. Contudo, verifica-se ainda que analisando a métrica MAC, o melhores desempenhos foram obtidos com as redes MLP e NARX, em um empate estatístico, seguido de RFC 2988, Holt-Winters, AES e MA.

Pondera-se ainda que o método MA, AES e HW apresentaram erros significativos quando da predição dos dados originais normalizados, obtendo valores elevados de MSE.

Verifica-se que a análise de métodos preditores de tráfego de redes unicamente através do parâmetro MSE pode ser complementada com a métrica MAC, visto que desta forma

tem-se uma representação mais fidedigna para a área de redes de computadores, pois esta métrica leva em consideração as amplitudes globalmente, o que no caso de tráfego de rede geralmente está associada com o tamanho dos pacotes.

Outra análise importante é no que concerne ao tempo de execução. Observando que todos os algoritmos foram executados no mesmo computador e sob iguais condições, a Tabela 21 apresenta o tempo, em segundos, necessário para dez rodadas de simulação dos algoritmos neurais atuando nos dados originais e nos contornos; o tempo de preparo dos dados, ou seja, o tempo de execução dos algoritmos extratores de contorno; e ao fim, demonstra o somatório da média (das redes MLP e NARX) com o tempo de preparo.

Tabela 21 - Tempo (em segundos) para execução de 10 rodadas dos algoritmos neurais

	MLP	NARX	Tempo de preparo (TP)	Tempo total = tempo médio (MLP e Narx) + TP
Dados originais	5,6718	5,9960	0	5,8339
Contornos por ECJF	0,7035	0,7998	0,186	0,9376
Contornos por ECJV	0,5302	0,6104	0,220	0,7903

Diante disso, verifica-se que o uso da abordagem de predição através de contornos apresenta uma redução de aproximadamente 87% do tempo necessário para finalizar a operação quando comparado com o mesmo procedimento realizado diretamente nos dados originais.

Percebe-se ainda que realizar predição de tráfego de redes a partir de seus contornos é bastante promissor, pois neste trabalho foi possível obter valores reduzidos de com uso de contornos; com destaque para o algoritmo proposto ECJV, obtendo o melhor desempenho na redução, 69,5%, e na capacidade de reduzir o valor MSE, bem como por ser o algoritmo mais rápido de ser processado.

4.5 Conclusão

Neste capítulo foram apresentados os resultados obtidos por meio dos métodos preditores descritos nesta dissertação. Tais métodos foram avaliados segundo o parâmetro estatístico MSE e a métrica proposta MAC.

A métrica MAC mostrou-se bastante útil na análise de predição de traços de tráfego de redes, visto que o parâmetro MSE muitas vezes revela um valor que por si só não permite

comparar adequadamente um mecanismo com outro. Isto se deve ao fato de o MSE não considerar as amplitudes (representativas do tamanho do pacote ou agrupamento) globalmente, apenas de forma pontual.

Dos mecanismos preditores, o que obteve melhor resultado quando analisando apenas o MSE foi o RFC 2988. Contudo, observou-se que, isoladamente, esta métrica não é um indicativo seguro para eleição de um mecanismo considerando se este obteve ou não bons resultados. A fim de preencher essa lacuna, uma análise em conjunto com MAC foi realizada, indicando que os melhores métodos preditores foram os métodos RFC 2988, MLP e NARX.

Diante da vantagem da predição de contornos, foi descrito o benefício em utilizar o algoritmo de extração de contorno proposto ECJV, reduzindo significativamente os dados e garantindo uma boa representatividade, bem como por ser o mais veloz na execução, garantindo eficiência e economia de recursos computacionais e ambientais.

O próximo capítulo conclui este manuscrito, resumindo as contribuições, conclusões e limitações identificadas no projeto de pesquisa desenvolvido e documentado nesta dissertação, além de se apresentar possíveis temas e sugestões para trabalhos futuros.

Capítulo 5

Considerações Finais

A predição de traços de redes requer técnicas que possibilitem uma estimação segura no contexto temporal e também no quantitativo do tamanho global dos pacotes. Este trabalho propôs métodos neurais para a predição de contornos de um tráfego de *backbone* Internet baseados em janelas fixas e variáveis.

Dentre os modelos de tráfego de rede hoje disponíveis, o Capítulo 2 apresentou o modelo LRD (*Long Range Dependence*) que influenciou significativamente o curso do trabalho quanto às técnicas de predição adotadas.

No Capítulo 3 foram exibidas as propostas desta dissertação, que podem ser resumidas em (i) algoritmo extrator de contornos com janelas fixa (ECJF) e variável (ECJV) e (ii) uso de redes neurais MLP e NARX na predição de contornos de traços de redes.

Os resultados foram amplamente expostos no Capítulo 4, informando-se ainda que foram analisados não apenas através da métrica MSE (*Mean Square Error*), mas também pela métrica desenvolvida neste trabalho a fim de considerar a predição no contexto do tamanho dos pacotes globalmente, o MAC (*Métrica das Amplitudes Conjugadas*).

As experimentações foram exitosas no que se refere à redução de dados de entrada dos preditores através dos algoritmos de extração de contornos, conseguindo manter a relação existente na série original e reduzindo drasticamente o tempo necessário para predição. As redes MLP e NARX, e o estatístico RFC 2988, forneceram bons resultados de predição, principalmente quando em atuação com os contornos obtidos por ECJF e ECJV.

5.1 Limitações do trabalho

As previsões em curto período resultantes dos métodos propostos é um dos aspectos que limita este trabalho; todavia, este aspecto é minimizado pela capacidade de redução dos algoritmos extratores de contorno propostos, pois, três pontos de contorno preditos significam, neste trabalho, aproximadamente dez pontos de dados do tráfego original.

Assim, esta dissertação se propõe a apresentar mecanismos de previsão neurais para atuação junto a contornos de tráfego aglomerado reais.

5.2 Trabalhos futuros

Os conhecimentos produzidos a partir deste trabalho representam o ponto de partida para a concepção de novas associações e experimentações, dentre as quais se podem recomendar: (i) aplicações para previsão de tráfego em redes móveis e sem fio; (ii) desenvolvimento de novas abordagens neuronais; (iii) previsões em longo prazo; e (iv) previsão de contornos de tráfego com base em indicadores de auto-similaridade e multifractalidade, e.g. parâmetro de *Hurst* e expoente de *Hölder*.

Como sugestão, recomenda-se experimentações de técnicas para a recuperação de um tráfego original a partir dos contornos preditos. Para tanto o uso de parâmetros estatísticos, de auto-similaridade e de multifractalidade tornar-se-ão imprescindíveis.

Como possibilidades para trabalhos futuros sugerem-se duas para o uso dos dados obtidos pelo neuropreditor: (i) provimento mais racional dos recursos de rede: assim, é possível contribuir na alocação de banda passante para acesso sem-fio (onde o espectro é limitado) e planejamento de capacidade; e (ii) detecção de anomalia de tráfego IP: considerando as constantes tentativas de invasões a redes e haja vista a alta variabilidade e irregularidade do tráfego da Internet há uma dificuldade intrínseca na distinção de fluxos normais de anômalos.

No tocante à aplicabilidade prática da nossa proposta, todos os algoritmos propostos poderiam ser embutidos em um roteador de borda de um provedor ISP. Deste modo, além do neuropreditor adaptar-se à sazonalidade do tráfego do *backbone*, posto que seu aprendizado dá-se de forma *online*, o referido provedor poderia propor a seus clientes contratos SLA (*Service Level Agreement*) personalizados, otimizando a alocação de recursos.

Referências bibliográficas

- Abry, P., & Veitch, D. (1998). Wavelet analysis of long-range-dependent traffic. *IEEE Trans. Info. Theory*, vol. Vol.44 , pp. 2-15.
- Allman, M., & Paxson, V. (1999). On estimating end-to-end network path properties. *Proceedings of ACM SIGCOMM* , pp. 263–274.
- Aragão Jr., J. B. (2008). *Computação Neural para Predição do Atraso de Execução de Pacotes de Voz em Redes IP. Dissertação (Mestrado em Engenharia de Teleinformática)*. Universidade Federal do Ceará.
- Aragão Jr., J. B., & Barreto, G. A. (2010). Novel approaches for online playout delay prediction in VoIP applications using time series models. *Comput Electr Eng (2010)*, doi:10.1016/j.compeleceng.2009.12.006 .
- Beran, J. (1994). *Statistics for long-memory processes*. Chapman & Hall.
- Bermúdez, J. D., Segura, J., & Vercher, E. (2005). Holt-Winters forecasting: an alternative formulation applied to UK air passenger data. *Technical Report of the Operations Research Center, Universidad Miguel Hernandez de Elche , N. I-2005-08*, pp. 1-16.
- Bianchi, G. R. (2005). Um modelo multifractal aplicado a predição de trafego de redes-proposta e avaliação. *Dissertação (Mestrado) - Mestrado em Engenharia Elétrica - Universidade Estadual de Campinas*.
- Brakmo, L. S., & Peterson, L. L. (1995). TCP Vegas: End to end congestion avoidance on global Internet. *IEEE Journal of Selected Areas in Communications*, vol. 13 , pp. 1465–1480.

- Carvalho, M. M. (1998). *Predição de Tráfego Auto-Similar em Redes de Faixa Larga*. Tese de Mestrado: Unicamp.
- Cipra, T. (1992). Robust exponential smoothing. *Journal of Forecast*. 11 , pp. 57-69.
- Cortez, P., Rio, M., Rocha, M., & Sousa, P. (2006). Internet Traffic Forecasting using Neural Networks. *International Joint Conference on Neural Networks - IJCNN 2006* , 2635-2642.
- Cortez, P., Rio, M., Sousa, P., & Rocha, M. (2007). Topology Aware Internet Traffic Forecasting Using Neural Networks. *Internet Corporation for Assigned Names and Numbers (2)*, pp. 445-454
- Croux, C., Gelper, S., & Fried, R. (2008). Computational aspects of robust Holt-Winters smoothing based on M-estimation. *Applications of Mathematics -PRAHA-, Vol. 53; N. 3* , pp. 163-176 .
- Crovella, M. E., & Bestavros, A. (1996). Self-similarity in world wide web traffic - evidence and possible causes. *Proceedings of ACM Sigmetrics* , 160-169.
- Diaconescu, E. (2008). The use of NARX neural networks to predict chaotic time series. *WSEAS Transactions Computer Research 3* , pp. 182-191.
- Gardner Jr., E. S. (2006). Exponential smoothing: The state of the art – Part II. *International Journal of Forecasting*, 22 , p. 637–666.
- Gardner Jr., E. S., & Diaz-Saiz, J. (2008). Exponential smoothing in the telecommunications data. *International Journal of Forecasting*, Volume 24, Issue 1 , pp. 170-174.
- Gomes, D. G., Agoulmine, N., Bennani, Y., & de Souza, J. N. (2007). Predictive connectionist approach for VoD bandwidth management. *Computer Communications* 30, 10 (Jul. 2007), 2236-2247. DOI=<http://dx.doi.org/10.1016/j.comcom.2007.05.005> .
- Haykin, S. (1994). *Neural Networks: A Comprehensive Foundation*. Englewood Cliffs, NJ: Macmillan Publishing Company.
- Karagiannis, T., Molle, M., & Faloutsos, M. (2004). Long-range dependence ten years of Internet traffic modeling. *Internet Computing, IEEE*, vol.8, no.5 , 57-64.
- Kleinrock, L. (2002). Creating a Mathematical Theory of Computer Networks. *INFORMS-Operations Research* 50, 1, (pp. 125-131).

- Kotsialos, A., Papageorgiou, M., & Poulimenos, A. (2005). Long-term sales forecasting using Holt-Winters and neural network methods. *Journal of Forecast* , pp. 353-368.
- Kurose, J. F., & Ross, K. W. (2006). *Redes de Computadores e a Internet: uma abordagem top-down*. Pearson Addison Wesley.
- Leland, W. E., & Wilson, D. V. (1991). High time-resolution measurement and analysis of LAN traffic: Implications for LAN interconnection. *INFOCOM (3) [S.l.: s.n.]*, (pp. 1360-1366).
- Leland, W. E., Taqqu, M. S., Willinger, W., & Wilson, D. V. (1994). On the self similar nature of Ethernet Traffic (extended version). *IEEE/ACM Transactions on Networking*, v.2, n.1 , pp. 1-15.
- Leontaritis, I. J., & Billings, S. A. (1985). Input-output parametric models for nonlinear systems - Part I: deterministic nonlinear systems. *International Journal of Control*, 41(2) , pp. 303–328.
- Li, Q., Zhao, X., Xu, M., & Liu, F. (2005). Network congestion prediction based on RFNN. *IEEE/SMC*, Vol. 3 , 2212-2217.
- Makridakis, S., Wheelwright, S., & Hyndman, R. J. (1998). *Forecasting: Methods and Applications*. Wiley.
- Mandelbrot, B. B., & Ness, J. V. (1968). Fractional brownian motions, fractional noises and applications. *SIAM Review*, vol. 10 , 422-437.
- Menezes Jr., J. M. P. (2006). *Redes neurais dinâmicas para predição e modelagem não-linear de séries temporais*. Dissertação (Mestrado) - Mestrado em Engenharia de Teleinformática, Universidade Federal do Ceará.
- Morettin, P. A., & Toloí, C. M. (2004). *Análise de Séries Temporais*. Edgard Blücher.
- Morettin, P. A., & Toloí, C. M. (1981). *Modelos para Previsão de Séries Temporais*. Edgard Blücher.
- Norgaard, M., Ravn, O., Poulsen, N. K., & Hansen, L. K. (2000). *Neural Networks for Modelling and Control of Dynamic Systems*. Springer.
- Orosz, G., Krauskopf, B., & Wilson, R. E. (2005). Bifurcations and multiple traffic jams in a car-following model with reaction-time delay. *Physica D* (211), 277-293.

- Owezarski, P., & Larrieu, N. (2004). Internet traffic characterization - an analysis of traffic oscillations. *7th IEEE International Conference on High Speed Networks and Multimedia Communications*, 96-107.
- Papagiannaki, D., Taft, N., Zhang, Z., & Diot, C. (2003). Long-term forecasting of Internet backbone traffic: Observations and initial models. *Proc. of IEEE Infocom* .
- Paxson, V., & Allman, M. (2000). *Computing Tcp's Retransmission Timer*. RFC Editor.
- Paxson, V., & Floyd, S. (1995). Wide area traffic: The failure of poisson modeling. *IEEE/ACM Transactions on Networking*, v.3, n.3 , 226-244.
- Roychoudhuri, L., & Al-Shaer, E. (2005). Real-Time Packet Loss Prediction based on End-to-end Delay Variation. *IEEE Transactions on Network and Service Management (TNSM)*, Volume: 2, Number: 1 , pp. 29-38.
- Scherrer, A., Larrieu, N., Borgnat, P., Owezarski, P., & Abry, P. (2006). Non-gaussian and long-memory statistical modeling of Internet traffic. *4th International Workshop on Internet Performance, Simulation, Monitoring and Measurements* .
- Shu, Y., Zhigang, J., Jidong, W., & Yang, O. W. (2000). Prediction-based admission control using FARIMA models. *International Conference on Communications - ICC 2000*, vol.3 , 1325-1329.
- Soysal, M., & Schmidt, E. (2010). Machine learning algorithms for accurate flow-based network traffic classification: Evaluation and comparison. *Performance Evaluation* . In Press, Corrected Proof, Available online 11 January 2010, ISSN 0166-5316, DOI: 10.1016/j.peva.2010.01.001.
- Vieira, F. H., & Ling, L. L. (2008). Adaptive Wavelet Based Multifractal Model Applied to the Effective Bandwidth Estimation of Network Traffic Flows. *IEE Proceedings. Communications*, v. 4 , 1-14.
- Wang, F., Li, D., & Zhao, Y. (2007). Prediction of Self-Similar Traffic and its Application in Network Bandwidth Allocation. *Wireless Communications, Networking and Mobile Computing*, (pp. 1980-1983). Shangai, China.
- Wheelwright, S. C., & Makridakis, S. (1985). *Forecasting Methods for Management*. New York: John Wiley & Sons Inc.
- Xie, Y.-B., Wang, W.-X., & Wang, B.-H. (2007). Modeling the coevolution of topology and traffic on weighted technological networks. *Physical Review E* , 026111.

Zhani, M. F., Elbiaze, H., & Kamoun, F. (2008). Analysis of prediction performance of training-based models using real network traffic. *Proc. of international symposium on performance evaluation of computer and telecommunication systems (SPECTS)* , 472-479.

Zhani, M. F., Elbiaze, H., & Kamoun, F. (2010). Analysis of prediction performance of training-based models using real network traffic. *International Journal of Computer Applications in Technology, Volume 37, Number 1* , pp. 10-19.

Apêndices

Apêndice A – Código-fonte do algoritmo ECJF

```

%% Tendencia Fixa - Janela = 3 (reformulado)
% v1 - Agosto de 2009 - Wendley Souza - wendley@ufc.br
% v2 - Novembro de 2009 - Wendley Souza - wendley@ufc.br

clear all;
close all;
load Caida-seg10.dat
temp=Caida_seg10;
temp=temp';
% clear Caida_seg1;

tic;

% Normalizacao dos dados de entrada [-1 , 1]
X=temp;
x = 2*(X-min(X))/(max(X)-min(X))-1;
temp=x;

soma_difer = 0;
diferencaTotal=0;
reducao=0;

tamanho = length(temp);

%% tendencias com máximos de grupos - janelas de tam. 3

janela=3; % tamanho utilizado na dissertação: 3
clear caidaTendMG;
soma_difer = 0;

for i=1:(fix(length(temp)/janela)) %fix: arredondamento p inteiro prox a
zero
    aux=max(temp(:, (i-1)*janela+1:(i-1)*janela+janela));
    caidaTendMG(i)=aux;
    soma_difer = soma_difer + sum(aux-temp(:, (i-1)*janela+1:(i-
1)*janela+janela));

```

```

end

reducao = (1-(length(caidaTendMG)/tamanho))*100;
diferencaTotal = soma_difer;

tempo=toc

reducao'
diferencaTotal'

temp=temp';
caidaTendMG=caidaTendMG';

save('ECJFs10.dat', 'caidaTendMG', '-ASCII');

%==== FIM DO ALGORITMO ECJF

```

Apêndice B – Código-fonte do algoritmo ECJV

```

%% Tendencias variaveis (reformulado)
% v1 - Agosto de 2009 - Wendley Souza - wendley@ufc.br
% v2 - Novembro de 2009 - Wendley Souza - wendley@ufc.br
% v3 - Janeiro de 2010 - Wendley Souza - wendley@ufc.br

clear all;
close all;
load Caida-seg10.dat
temp=Caida_seg10;
temp=temp';

soma_difer = 0;
diferencaTotal=0;
tamanho = length(temp);

tic;

% Normalizacao dos dados de entrada [-1 , 1]
X=temp;
x = 2*(X-min(X))/(max(X)-min(X))-1;
temp=x;

minimo=min(temp);
maximo=max(temp);
diferenca_maxima=maximo-minimo;

%% tendencias com máximos de grupos - janela variável

reducaoGlobal = 0;

```

```

diferencaGlobal = 0;
tamJAGGlobal = 0;
countGlobal = 0;

criterio=0.1; % utilizado na dissertacao 0.1

clear TendVar;

aux=temp(:,1);
TendVar(1)=aux;
count=1;

    for i=1:tamanho-1

        if soma_difer < criterio*diferenca_maxima
            indiceMin = i;
            if temp(:,i+1) > temp(:,i)
                aux=temp(:,i+1);
                TendVar(count)=aux;
                indiceMax=i+1;
                soma_difer = sum(aux-temp(:,indiceMin:indiceMax));
            end

            else
                indiceMax = i+1;
                diferencaTotal=diferencaTotal+soma_difer;

                count=count+1;
                soma_difer = 0;
            end;
        end;

        countGlobal=countGlobal+1;

    diferencaGlobal(countGlobal)=(1-(length(TendVar)/tamanho))*100;
    diferencaGlobal(countGlobal)=diferencaTotal;
    tamJAGGlobal(countGlobal)=tamanho/length(TendVar);

    tempo = toc

    diferencaGlobal'
    diferencaGlobal'
    tamJAGGlobal'

    TendVar=TendVar';

    save('ECJVs10.dat', 'TendVar', '-ASCII')

%==== FIM DO ALGORITMO ECJV

```

Apêndice C – Código-fonte do algoritmo MLP

```

% Predição usando MLP com aprendizado On-line - Wendley - wendley@ufc.br
% Baseado no algoritmo do prof. Guilherme Barreto -
% http://deti.ufc.br/~guilherme/
% Fevereiro de 2010

clear all; close all; clc;

% Carrega DADOS
%=====

for s=1:10

% load CNorms1a10.mat %--> Dados originais normalizados
% trace = CNorms1a10;

% load ECJFs1a10.mat %--> Dados de contornos obtidos por ECJF
% trace = ECJFs1a10;

load ECJVs1a10.mat %--> Dados de contornos obtidos por ECJV
trace = ECJVs1a10;

countG=1;

trace = trace(:,s);

if s==1 tamanho = 283
elseif s==2 tamanho = 297
elseif s==3 tamanho = 294
elseif s==4 tamanho = 271
elseif s==5 tamanho = 280
elseif s==6 tamanho = 274
elseif s==7 tamanho = 294
elseif s==8 tamanho = 291
elseif s==9 tamanho = 270
elseif s==10 tamanho = 276
end; % Descomentar se usar ECJV

% tamanho = length(trace); % Comentar se usar ECJV

clear alvos;
clear matriz_entradas;

% Normalizacao dos dados de entrada [0 , 1]
X=trace;
x = (X-min(X))/(max(X)-min(X));
trace=x;

```



```

vetmax= 1;

for tt=1:cP
    EQ=0;
    vmax= max(dados(:,tt));
    if vmax>vetmax; vetmax=vmax; end

    % CAMADA OCULTA
    entradas= dados(:,tt)/max(dados(:,tt));
    X= [-1; entradas]; % Constroi vetor de entrada com adicao da
    entrada x0=-1 (Bias)
    Ui= WW * X; % Ativacao (net) dos neuronios da camada oculta
    Yi= 1./(1+exp(Ui)); % Saida entre [0,1] (funcao logistica)
    Yi_out(:,tt)= Yi;
    % CAMADA DE SAIDA
    Y= [-1; Yi]; % Constroi vetor de entrada DESTA CAMADA com adicao
    da entrada y0=-1 (Bias)
    Uk= MM * Y; % Ativacao (net) dos neuronios da camada de saida
    Ok(tt)= (1./(1+exp(-Uk))) * max(dados(:,tt)); % Saida não linear
    % Ok(tt) = Uk; %Saida linear f(x)=x;
    matriz_entradas(7,tt)= Ok(tt); % gravando a saída da rede na
    matriz_entradas para uma melhor visualização.
    % CALCULO DO ERRO
    Ek= alvos(tt) - Ok(tt); % erro entre a saida desejada e a saida da
    rede
    Erro(tt)= Ek;
    matriz_entradas(8,tt)= Ek; % gravando o erro de estimacao da rede na
    matriz_entradas para uma melhor visualização.

    % CALCULO DOS GRADIENTES LOCAIS
    %Dk= Ok.*(1 - Ok); % derivada da sigmoide logistica (camada de
    saida)
    Dk= 1; % derivada da funcao identidade utilizada na saida.
    DDk= Ek.*Dk; % gradiente local (camada de saida)

    Di = Yi.*(1 - Yi); % derivada da sigmoide logistica
    (camada oculta)
    DDi= Di.*(MM(:,2:end)'+DDk); % gradiente local (camada oculta)

    % AJUSTE DOS PESOS: camada oculta -> camada de saida
    MM_aux= MM;
    MM= MM + eta*DDk*Y' + mom*(MM - MM_old);
    MM_old= MM_aux;

    % AJUSTE DOS PESOS: entrada -> camada oculta
    WW_aux= WW;
    WW= WW + eta*DDi*X' + mom*(WW - WW_old);
    WW_old= WW_aux;

end

clear Yi;
clear X;
clear Ui;

```

```

clear Yi_out;
clear WW;
clear entradas;
clear xx;

% METRICAS

% ===== MAC - Metrica das Amplitudes Conjugadas

soma=0;
somaP=0;

soma = sum(alvos);
somaP = sum(Ok);

s1 = soma/length(alvos);
s2 = somaP/length(Ok);

MAC = s2/s1;

% ===== MSE - Erro Quadrático Médio

MSE= (sum(Erro.^2))/length(Erro); %MSE

tempo(count)=toc;

MACTt(count) = MAC;
MSEtt(count) = MSE;

% clc
% andamento=countG/1.55

count = count+1;

% countG=countG+1;
s
end

tempott(s,:)=tempo;

MACTtt(s,:) = MACTt';
MSEttt(s,:) = MSEtt';

clear Ok;

end

```

```

MACttt=MACttt';
MACttt=1-MACttt;

MSEttt=MSEttt';
tempott=tempott';

mediaMAC = mean(MACttt)'
mediaMSE = mean(MSEttt)'
tempoTT = sum(tempott)

relacao = (tempoTT + 2*mediaMAC + 3*mediaMSE)/6;

relacao
min(relacao)

conj(:,1)=mediaMAC;
conj(:,2)=mediaMSE;
conj(:,3)=tempoTT;

conj3=conj;

save('conj3.mat', 'conj3');

%% Graficos de determinações (de mérito)

% xx=0:0.05:0.6;
%
% % createfigure6(conj); % Gera grafico da definição de nr. neuronios na
camada escondida
% subplot(3,1,1)
% plot(xx,mediaMSE,'ks-')
% title('MSE médio das 10 rodadas')
% % xlabel('Nr. de neurônios na camada escondida')
% % xlabel('Taxa de aprendizagem')
% xlabel('Valor do termo de momento')
% ylabel('MSE')
%
% subplot(3,1,2)
% plot(xx,mediaMAC,'ko-')
% title('MAC médio das 10 rodadas')
% % xlabel('Nr. de neurônios na camada escondida')
% % xlabel('Taxa de aprendizagem')
% xlabel('Valor do termo de momento')
% ylabel('MAC')
%
% subplot(3,1,3)
% plot(xx,tempoTT,'k*-')
% title('Tempo em segundos das 10 rodadas')
% % xlabel('Nr. de neurônios na camada escondida')
% % xlabel('Taxa de aprendizagem')
% xlabel('Valor do termo de momento')

```

```
% ylabel('Tempo (s)')
%
% % Fim do código de Graficos de determinações (de mérito)
```

```
%==== FIM DO ALGORITMO MLP
```

Apêndice D – Código-fonte do algoritmo NARX

```
% Predição usando NARX com aprendizado On-line - Wendley - wendley@ufc.br
% Baseado no algoritmo do prof. Guilherme Barreto -
% http://deti.ufc.br/~guilherme/
% Fevereiro de 2010
```

```
clear all; close all; clc;
```

```
% Carrega DADOS
%=====
```

```
for s=1:10
```

```
load CNorms1a10.mat %--> Dados originais normalizados
trace = CNorms1a10;
```

```
% load ECJFs1a10.mat %--> Dados de contornos obtidos por ECJF
% trace = ECJFs1a10;
```

```
% load ECJVs1a10.mat %--> Dados de contornos obtidos por ECJV
% trace = ECJVs1a10;
```

```
countG=1;
```

```
trace= trace(:,s);
```

```
% if s==1 tamanho = 283
% elseif s==2 tamanho = 297
% elseif s==3 tamanho = 294
% elseif s==4 tamanho = 271
% elseif s==5 tamanho = 280
% elseif s==6 tamanho = 274
% elseif s==7 tamanho = 294
% elseif s==8 tamanho = 291
% elseif s==9 tamanho = 270
% elseif s==10 tamanho = 276
```



```

% DEFINE ARQUITETURA DA REDE
%=====
Nh= 54; % No. de neuronios na camada oculta
No= 1; % No. de neuronios na camada de saida

eta= 0.05 % Passo de aprendizagem
mom= xx; % Fator de momento

% Inicia matrizes de pesos
WW= 0.01*rand(Nh,1P+6); % Pesos entrada -> camada oculta
WW_old= WW; % Necessario para termo de momento

MM= 0.01*rand(No,Nh+1); % Pesos camada oculta -> camada de saida
MM_old= MM; % Necessario para termo de momento
vetmax= 1;

Ok = [1 1 1 1 1];

for tt=1:(cP)
    EQ=0;
    vmax= max(dados(:,tt));
    if vmax>vetmax; vetmax=vmax; end

    % CAMADA OCULTA
    entradas(1:1P,:) = dados(:,tt)/max(dados(:,tt));
    entradas(1P+1:1P+5,:) = Ok(tt:tt+4); % Retroalimentação
    X= [-1; entradas]; % Constroi vetor de entrada com adicao da
entrada x0=-1 (Bias)
    Ui= WW * X; % Ativacao (net) dos neuronios da camada
oculta
    Yi= 1./(1+exp(Ui)); % Saida entre [0,1] (funcao logistica)
    Yi_out(:,tt)= Yi;
    % CAMADA DE SAIDA
    Y= [-1; Yi]; % Constroi vetor de entrada DESTA CAMADA com
adicao da entrada y0=-1 (Bias)
    Uk= MM * Y; % Ativacao (net) dos neuronios da camada de
saida
    Ok(tt+5)= (1./(1+exp(-Uk))) * max(dados(:,tt)); % Saida não
linear
    %Ok(tt+5) = Uk; %Saida linear f(x)=x;
    matriz_entradas(7,tt)= Ok(tt+1); % gravando a saida da rede na
matriz_entradas para uma melhor visualização.
    % CALCULO DO ERRO
    Ek= alvos(tt) - Ok(tt+5); % erro entre a saida desejada e a
saida da rede
    Erro(tt)= Ek;
    matriz_entradas(8,tt)= Ek; % gravando o erro de estimacao da rede
na matriz_entradas para uma melhor visualização.

    % CALCULO DOS GRADIENTES LOCAIS
    %Dk= Ok.*(1 - Ok); % derivada da sigmoide logistica (camada de
saida)

```

```

Dk= 1; % derivada da função identidade utilizada na saída.
DDk= Ek.*Dk; % gradiente local (camada de saída)

Di = Yi.*(1 - Yi); % derivada da sigmoide logistica
(camada oculta)
DDi= Di.*(MM(:,2:end)'.*DDk); % gradiente local (camada oculta)

% AJUSTE DOS PESOS: camada oculta -> camada de saída
MM_aux= MM;
MM= MM + eta*DDk*Y' + mom*(MM - MM_old);
MM_old= MM_aux;

% AJUSTE DOS PESOS: entrada -> camada oculta
WW_aux= WW;
WW= WW + eta*DDi*X' + mom*(WW - WW_old);
WW_old= WW_aux;
end

clear Yi;
clear X;
clear Ui;
clear Yi_out;
clear WW;
clear entradas;
clear xx;

% METRICAS

% ===== MAC - Metrica das Amplitudes Conjugadas

soma=0;
somaP=0;

soma = sum(alvos);
somaP = sum(Ok);

s1 = soma/length(alvos);
s2 = somaP/length(Ok);

MAC = s2/s1;

% ===== MSE - Erro Quadrático Médio

MSE= (sum(Erro.^2))/length(Erro); %MSE

tempo(count)=toc;

MACTt(count) = MAC;
MSEtt(count) = MSE;

```

```

clc
andamento=countG/1.6

count = count+1;

countG=countG+1;

end

tempott(s,:)=tempo;

MACTtt(s,:) = MACTt';
MSEttt(s,:) = MSEtt';

% clear Ok;

end

MACTtt=MACTtt';
MACTtt=1-MACTtt;

MSEttt=MSEttt';
tempott=tempott';

mediaMAC = mean(MACTtt')'
mediaMSE = mean(MSEttt')'
tempoTT = sum(tempott')'

relacao = (tempoTT + 2*mediaMAC + 3*mediaMSE)/6; % para definição

relacao
min(relacao)

conj(:,1)=mediaMAC;
conj(:,2)=mediaMSE;
conj(:,3)=tempoTT;

%% Graficos de determinações (de mérito)

% xx=0:0.05:0.6;
%
% % createfigure6(conj); % Gera grafico da definição de nr. neuronios na
camada escondida
% subplot(3,1,1)
% plot(xx,mediaMSE,'ks-')
% title('MSE médio das 10 rodadas')
% % xlabel('Nr. de neurônios na camada escondida')
% % xlabel('Taxa de aprendizagem')
% xlabel('Valor do termo de momento')

```

```
% ylabel('MSE')
%
% subplot(3,1,2)
% plot(xx,mediaMAC,'ko-')
% title('MAC médio das 10 rodadas')
% % xlabel('Nr. de neurônios na camada escondida')
% % xlabel('Taxa de aprendizagem')
% xlabel('Valor do termo de momento')
% ylabel('MAC')
%
% subplot(3,1,3)
% plot(xx,tempoTT,'k*-')
% title('Tempo em segundos das 10 rodadas')
% % xlabel('Nr. de neurônios na camada escondida')
% % xlabel('Taxa de aprendizagem')
% xlabel('Valor do termo de momento')
% ylabel('Tempo (s)')
%
% % Fim do código de Graficos de determinações (de mérito)

%==== FIM DO ALGORITMO NARX
```

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)