

YVO MARCELO CHIARADIA MASSELLI

**SUGESTÃO DE UM SISTEMA HÍBRIDO NEURO-GENÉTICO COMO
ALTERNATIVA À AVALIAÇÃO DE CURVAS DE CONSUMO DE
ENERGIA ELÉTRICA**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Energia da UNIFEI como parte dos requisitos para a obtenção do título de Mestre em Engenharia da Energia.

Orientador: Germano Lambert Torres
Co-Orientador: Luiz Eduardo Borges da Silva

Departamento de Engenharia da Energia

Universidade Federal de Itajubá

Março de 2005

Dedicatória

à minha família

Agradecimentos

- Ao Amigo e Professor Germano Lambert Torres;
- Ao apoio da concedido pela CAPES/CNPq;
- Aos colegas;
- Àqueles que mesmo distantes olham por mim;

RESUMO

A dissertação a seguir aborda os conceitos, aplicações e hibridização de duas técnicas de Inteligência Artificial: redes neurais artificiais e algoritmos genéticos.

Em primeiro plano são apresentados os conceitos de redes neurais artificiais, suas principais características, vantagens e aplicações práticas.

Em seguida são expostos conceitos da técnica conhecida por algoritmos genéticos, assim como suas características mais importantes, seu embasamento biológico, como é desenvolvida e onde é aplicada.

A partir do conhecimento das técnicas anteriores parte-se para a associação entre elas, constituindo assim um sistema híbrido, onde a cooperação entre redes neurais artificiais e algoritmos genéticos passa a ser o ponto de foco do presente trabalho.

Apresentados os sistemas híbridos, uma aplicação prática é desenvolvida no intuito de comparar os resultados apresentados por este novo modelo, onde há a hibridização das técnicas acima citadas, com os gerados pela rede neural artificial atuando isoladamente.

Esta aplicação se baseia na avaliação de curvas de consumo de energia elétrica (curvas de duração de carga ou simplesmente curvas de carga), o que proporciona uma alternativa à gestão do sistema de distribuição. Para a análise desenvolvida foram escolhidos os três principais setores consumidores de energia elétrica, o residencial, comercial e o industrial.

O sistema desenvolvido neste trabalho tem como objetivo a identificação do consumo relativo a cada um destes setores mediante a apresentação de uma curva de consumo adquirida em um ponto qualquer do sistema de distribuição. O conhecimento destes valores proporciona uma série de vantagens que serão descritas adiante e permite o desenvolvimento de uma poderosa ferramenta para a gestão do sistema de distribuição de energia elétrica.

ABSTRACT

The following dissertation approach the concepts, applications and hybridization of two Artificial Intelligent techniques: artificial neural networks and genetic algorithms.

Firstly the concepts, main features, advantages and a practical application of artificial neural networks are introduced.

Next, the presentation of technic concepts well know by genetic algorithm, as well your main features, your biological environment, how is developed and where is applied.

From the early techniques knowledge, start the associations between them, building a hybrid system, where the cooperation between artificial neural networks and genetic algorithms turn into focus of this work.

Showed the hybrid system, a practical application is developed to results compare on this new model, where exist the hybridization of below techniques quoted with those generated from neural network isolate working.

This application is based on the assessment of electrical energy load curves (or Load duration curve) that give a alternative to the distribution system management.

To the developed analyses was choose three main electrical energy consumers sectors: residential, commercial and industrial.

The developed system on this work aim to the relative consumption identification to each one of these sectors on the face of a load curve presentation acquired on any point of distribution system. These values knowledge give serie advantages that will be describes and allow the development of a powerful tool for the electrical energy distribution system management.

LISTA DE ILUSTRAÇÕES	VIII
LISTA DE TABELAS.....	XI
LISTA DE SIGLAS.....	1
1 – INTRODUÇÃO.....	2
1.1 MOTIVAÇÃO.....	2
1.2 OBJETIVOS DO TRABALHO.....	2
1.3 DESCRIÇÃO DO TRABALHO.....	3
1.4 ORGANIZAÇÃO DA DISSERTAÇÃO.....	5
2 - REDES NEURAIS ARTIFICIAIS	7
2.1 INTRODUÇÃO.....	7
2.2 HISTÓRICO DAS REDES NEURAIS ARTIFICIAIS.....	7
2.3 O NEURÔNIO ARTIFICIAL.....	11
2.4 PERCEPTRON.....	12
2.5 FUNÇÃO DE ATIVAÇÃO.....	13
2.6 TOPOLOGIA.....	14
2.7 APRENDIZADO.....	16
2.7.1 Aprendizado Supervisionado.....	16
2.7.2 Aprendizado Não Supervisionado.....	17
2.8 REDES COM MÚLTIPLAS CAMADAS.....	17
2.9 BACKPROPAGATION.....	19
2.9.1 Considerações importantes.....	19
3 - ALGORITMOS GENÉTICOS	21
3.1 INTRODUÇÃO.....	21
3.2 INSPIRAÇÃO BIOLÓGICA.....	21
3.2 O ALGORITMO GENÉTICO.....	22
3.3.1 População Inicial.....	24
3.3.2 Função Aptidão.....	26
3.3.3 Operadores Genéticos.....	28
3.3.3.1 Crossover.....	29
3.3.3.2 Mutaç�o.....	30
3.3.4 Crit�rios de Parada.....	31
3.3.5 Par�metros Importantes.....	32
3.3.5.1 N�mero de Indiv�duos.....	32

3.3.5.2 Taxa de Crossover	33
3.3.5.2 Taxa de Mutação	33
4 - SISTEMAS HÍBRIDOS	34
4.1 INTRODUÇÃO	34
4.2 SISTEMA HÍBRIDO PROPOSTO	34
4.2.1 Considerações Iniciais	34
4.2.2 O Sistema Híbrido.....	35
4.2.3 Etapas do Processo	37
4.3 VANTAGENS DO HIBRIDISMO.....	39
4.4 RESULTADOS ESPERADOS	39
5 - ESTUDO DE CASO	41
4.5 INTRODUÇÃO	41
5.2 APLICAÇÃO PRÁTICA.....	41
5.3 DEFINIÇÃO DOS PADRÕES DE TREINAMENTO	44
5.4 SIMULAÇÕES ENVOLVENDO PADRÕES NATURAIS SIMPLES E COMPOSTOS.....	45
5.4.1 Mudanças no Conjunto de Treinamento	52
5.5 DEFINIÇÃO DOS NOVOS PADRÕES DE TREINAMENTO	59
5.5.1 Simulações envolvendo composições entre os novos padrões	65
5.6 CONSIDERAÇÕES SOBRE OSCILAÇÕES	71
6 - CONCLUSÕES E TRABALHOS FUTUROS.....	80
7 - REFERÊNCIAS BIBLIOGRÁFICAS.....	83

LISTA DE ILUSTRAÇÕES

Figura 2.1: Modelo de um neurônio artificial proposto por McCulloch e Pitts.....	11
Figura 2.2: Modelo do Perceptron proposto por Rosenblatt	12
Figura 2.3: Representação Básica de uma rede MLP	15
Figura 2.4: Representação Básica de uma rede MLP Recorrente	15
Figura 2.5: Ilustração de uma superfície qualquer e seus vetores gradientes	18
Figura 3.1: Fluxograma Básico dos Algoritmos Genéticos	24
Figura 3.2: Seleção pelo Método da Roleta.....	27
Figura 3.3: Exemplo de crossover para a codificação binária	29
Figura 3.4: Exemplo de crossover para a codificação real	30
Figura 3.5: Exemplo de mutação para a codificação binária	31
Figura 3.6: Exemplo de mutação para a codificação real.....	31
Figura 4.1: Treinamento da RNA através de métodos baseados em Gradiente Descendente	36
Figura 4.2: Treinamento da RNA através de Algoritmos Genéticos	37
Figura 4.3: Exemplo de uma Curva Multimodal	40
Figura 5.1: Participação dos Principais Setores de Atividade no consumo de energia elétrica no país (em porcentagem).....	42
Figura 5.2: Aplicação Prática.....	43
Figura 5.3: Padrões Naturais.....	44
Figura 5.4: Conjunto de treinamento composto por impulsos e pulsos retangulares.....	46
Figura 5.5: Conjunto de validação formado pela composição de impulsos e pulsos retangulares ...	46
Figura 5.6: Conjunto de treinamento formado por parábolas e impulsos.....	48
Figura 5.7: Conjunto de validação formado pela composição de parábolas e impulsos.....	48
Figura 5.8: Conjunto de treinamento formado pela composição de parábolas e ondas triangulares	49
Figura 5.9: Conjunto de validação formado pela composição de parábolas e ondas triangulares ...	49
Figura 5.10: Conjunto de treinamento formado por ondas triangulares e pulsos retangulares.....	50
Figura 5.11: Conjunto de validação formado pela composição de ondas triangularese pulsos retangulares.....	50
Figura 5.12: Conjunto de treinamento formado por impulsos e pulsos retangulares, assim como composições destes.....	53
Figura 5.13: Conjunto de validação formado por composições de impulsos e pulsos retangulares...	53
Figura 5.14: Conjunto de treinamento formado por parábolas e impulsos, assim como composições destes.....	55
Figura 5.15: Conjunto de validação formado por composições de parábolas e impulsos.....	55
Figura 5.16: Conjunto de treinamento formado por parábolas e ondas triangulares, assim como composições destes.....	56
Figura 5.17: Conjunto de validação formado por composições de parábolas e ondas triangulares ...	56

Figura 5.18: Conjunto de treinamento formado por ondas triangulares e pulsos retangulares, assim como composições destes.....	58
Figura 5.19: Conjunto de validação formado por composições de ondas triangulares e pulsos retangulares.....	58
Figura 5.20: Curvas adotadas como referências aos Setores Residencial, Comercial e Industrial	60
Figura 5.21: Conjunto de treinamento formado por curvas representativas do setor residencial, com várias amplitudes.....	61
Figura 5.22: Conjunto de validação formado por curvas do setor residencial, com amplitudes diferentes das utilizadas para o treinamento	61
Figura 5.23: Conjunto de treinamento formado por curvas representativas do setor industrial, com várias amplitudes.....	62
Figura 5.24: Conjunto de validação formado por curvas do setor industrial, com amplitudes diferentes das utilizadas para o treinamento.....	62
Figura 5.25: Conjunto de treinamento formado por curvas representativas do setor comercial, com várias amplitudes.....	63
Figura 5.26: Conjunto de validação formado por curvas do setor comercial, com amplitudes diferentes das utilizadas para o treinamento.....	63
Figura 5.27: Conjunto de treinamento formado por curvas representativas dos setores residencial e comercial, assim como composições das mesmas	66
Figura 5.28: Conjunto de validação formado por composições de curvas representativas dos setores residencial e comercial.....	66
Figura 5.29: Conjunto de treinamento formado por curvas representativas dos setores residencial e industrial, assim como composições das mesmas	68
Figura 5.30: Conjunto de validação formado por composições de curvas representativas dos setores residencial e industrial.....	68
Figura 5.31: Conjunto de treinamento formado por curvas representativas dos setores industrial e comercial, assim como composições das mesmas	69
Figura 5.32: Conjunto de validação formado por composições de curvas representativas dos setores industrial e comercial.....	69
Figura 5.33: Conjunto de treinamento formado por curvas representativas dos setores residencial e comercial, assim como composições das mesmas submetidas a oscilações	72
Figura 5.34: Conjunto de validação formado por composições de curvas representativas dos setores industrial e comercial submetidas a oscilações	73
Figura 5.35: Conjunto de treinamento formado por curvas representativas dos setores industrial e comercial, assim como composições das mesmas submetidas a oscilações.....	74
Figura 5.36: Conjunto de validação formado por composições de curvas representativas dos setores industrial e comercial submetidas a oscilações	75
Figura 5.37: Conjunto de treinamento formado por curvas representativas dos setores residencial e industrial, assim como composições das mesmas submetidas a oscilações.....	77

Figura 5.38: Conjunto de validação formado por composições de curvas representativas dos setores residencial e industrial submetidas a oscilações 77

LISTA DE TABELAS

Tabela 3.1: Exemplos de Codificações.....	25
Tabela 5.1: Resultados da simulação para padrões como impulsos e pulsos retangulares	47
Tabela 5.2: Resultados da simulação para padrões como parábolas e impulsos	48
Tabela 5.3: Resultados da simulação para padrões como parábolas e ondas triangulares	50
Tabela 5.4: Resultados da simulação para padrões como ondas triangulares e pulsos	
retangulares.....	51
Tabela 5.5: Parâmetros das RNA.....	52
Tabela 5.6: Resultados da simulação para padrões como impulsos e pulsos retangulares	54
Tabela 5.7: Resultados da simulação para padrões como parábolas e impulsos	55
Tabela 5.8: Resultados da simulação para padrões como parábolas e ondas triangulares	57
Tabela 5.9: Resultados da simulação para padrões como ondas triangulares e pulsos	
retangulares.....	58
Tabela 5.10: Resultados da simulação realizada com o padrão residencial	62
Tabela 5.11: Resultados da simulação realizada com o padrão industrial	63
Tabela 5.12: Resultados da simulação realizada com o padrão comercial	64
Tabela 5.13: Resultados da simulação realizada com os padrões residencial e comercial	67
Tabela 5.14: Resultados da simulação realizada com os padrões residencial e industrial	69
Tabela 5.15: Resultados da simulação realizada com os padrões industrial e comercial	70
Tabela 5.16: Resultados da simulação realizada com 10% de oscilações sobre as amplitudes dos padrões residencial e comercial.....	73
Tabela 5.17: Resultados da simulação realizada com 5% de oscilações sobre as amplitudes dos padrões residencial e comercial.....	74
Tabela 5.18: Resultados da simulação realizada com 10% de oscilações sobre as amplitudes dos padrões industrial e comercial.....	75
Tabela 5.19: Resultados da simulação realizada com 5% de oscilações sobre as amplitudes dos padrões industrial e comercial.....	76
Tabela 5.20: Resultados da simulação realizada com 10% de oscilações sobre as amplitudes dos padrões residencial e industrial.....	78
Tabela 5.21: Resultados da simulação realizada com 5% de oscilações sobre as amplitudes dos padrões residencial e industrial.....	78

LISTA DE SIGLAS

RNA: Rede Neural Artificial

AG: Algoritmo Genético

BP: Backpropagation

MLP: Multi-layer Perceptron

SSE: Sum Squared Error

MSE: Mean Squared Error

XOR: função Ou-Exclusivo

BL α X: Crossover Média

KWh: Kilo-Watt-hora

KVA: Kilo-Volt-Ampere

CAPÍTULO I

1 – INTRODUÇÃO

1.1 MOTIVAÇÃO

Com o passar do tempo surge, de forma quase que inerente ao processo evolutivo, uma “interdependência” e o conseqüente aumento de intimidade na relação homem-máquina. Visto que o lado “máquina” é a cada dia mais provido de conhecimento, o lado “homem” sente vorazmente o desejo de “máquinas” cada vez mais inteligentes e que, de certa forma, o substitua em tarefas das mais diversas naturezas.

Para que isso seja possível é necessário que, de alguma forma, o homem consiga embutir nas máquinas o seu conhecimento adquirido ao longo do tempo. Nesta direção caminha a ciência, mais especificamente àquela denominada Inteligência Artificial, que tem como propósito desenvolver técnicas que permitam reproduzir em máquinas as características mais importantes do homem como pensar, tomar decisões e agir.

Diferenciadas por áreas de aplicação, diversas técnicas no ramo de Inteligência Artificial foram desenvolvidas na intenção otimizar e automatizar processos através da agregação de algum tipo de conhecimento às máquinas.

Este trabalho investiga técnicas de Redes Neurais Artificiais e Algoritmos Genéticos aplicadas ao setor elétrico. Primeiramente estas são analisadas separadamente e depois é sugerida a hibridização entre elas, no intuito de otimizar a aplicação prática proposta.

O detalhamento das técnicas e o desenvolvimento da aplicação proposta são apresentados ao longo desta dissertação.

1.2 OBJETIVOS DO TRABALHO

Os objetivos principais deste trabalho consistem em:

- Realizar uma análise detalhada sobre Redes Neurais Artificiais (RNA) e Algoritmos Genéticos (AG), de acordo com a aplicação proposta.
- Desenvolver um sistema híbrido, através da associação entre RNA e AG, que será proposto como alternativa aos tradicionais sistemas baseados em métodos numéricos para o treinamento de RNA.
- Avaliar o desempenho do sistema desenvolvido através da aplicação prática proposta. Tal sistema deverá ser capaz de, ante a apresentação de uma Curva de Carga qualquer, quantificar os níveis de consumo de energia elétrica relativos aos tipos de consumidores aos quais a curva se refere, assim como apontar variações nos patamares de consumo (Pico da Curva de Carga).
- Proporcionar, através da análise desenvolvida, a implementação de uma ferramenta adicional à gestão da rede elétrica de distribuição, como alternativa aos atuais modelos baseados em cálculos estatísticos.

1.3 DESCRIÇÃO DO TRABALHO

Esta dissertação trata em primeiro plano a técnica conhecida por Redes Neurais Artificiais (RNA). São introduzidos seus principais conceitos e em seguida, dados detalhes sobre as bases biológicas que inspiraram seu desenvolvimento, desde o primeiro modelo de neurônio artificial até o conhecido por perceptron que até hoje é utilizado. Também são expostos conceitos sobre funções de ativação, topologia, aprendizado e suas principais particularidades. Questões importantes, como o número de camadas intermediárias e a quantidade de neurônios necessária a cada uma delas, são também discutidas. Em seguida são conceituadas as redes de múltiplas camadas e o algoritmo utilizado para o treinamento destas.

Abordadas as principais questões sobre RNA o trabalho continua com o estudo da técnica conhecida por Algoritmos Genéticos (AG). São também apresentadas as aplicações práticas dos AG, seu embasamento biológico e as vantagens do emprego da técnica.

Dentro deste assunto são tratadas questões fundamentais ao processo de execução do algoritmo, como o alfabeto de codificação utilizado, a função de avaliação e os métodos de seleção. A seguir é discutida a importância da aplicação dos operadores genéticos: crossover e mutação, assim como as particulares de cada um deles.

O passo seguinte é a proposta de hibridização das técnicas apresentadas anteriormente.

Neste ponto é sugerido um sistema híbrido neuro-genético como alternativa ao sistema puramente neural. Então é realizada a aplicação prática do sistema proposto através da avaliação de curvas de consumo de energia elétrica. São consideradas as curvas típicas dos setores residencial, industrial e comercial e, a partir destas, o desafio será o de identificar e quantificar a porção referente a cada um destes setores em uma curva de carga extraída de um ponto qualquer da rede de distribuição. Para isto são desenvolvidos um sistema puramente neural e um sistema neuro-genético onde o treinamento das RNA é realizado pelo AG.

Inúmeras simulações foram realizadas com os dois sistemas e a análise dos resultados obtidos em cada uma delas são apresentados.

A seguir estão descritas as conclusões referentes à viabilidade dos sistemas propostos, assim como um comparativo entre os métodos utilizados para o treinamento das Redes Neurais Artificiais.

Finalmente apresentam-se as conclusões referentes ao trabalho, em especial sobre estudo de caso. Posteriormente são dadas sugestões de estudos futuros a partir do trabalho aqui apresentado.

1.4 ORGANIZAÇÃO DA DISSERTAÇÃO

Esta dissertação está dividida em seis capítulos adicionais, descritos a seguir:

O capítulo 2 apresenta a técnica conhecida por redes neurais artificiais. Primeiramente é colocada uma breve introdução assim como seu histórico. Depois são detalhados os fundamentos, particularidades e aplicações da mesma. O capítulo é encerrado com a apresentação de um dos pontos mais importantes sobre o assunto, os processos de aprendizagem e treinamento.

O capítulo 3 detalha a técnica de otimização denominada algoritmos genéticos. Neste capítulo são dadas as informações necessárias ao seu desenvolvimento e implementação. Pontos como fundamentos biológicos, ajustes de parâmetros e todos os processos envolvidos são expostos.

O capítulo 4 traz as informações necessárias à composição das técnicas anteriormente abordadas, constituindo assim um sistema híbrido. Este sistema será proposto como uma alternativa ao uso das redes neurais artificiais convencionais, onde o treinamento é realizado através do algoritmo conhecido por Backpropagation. As vantagens do uso deste tipo de sistema são também comentadas.

O capítulo 5 traz uma aplicação prática do sistema híbrido proposto. Esta aplicação possibilitará a avaliação do método convencional de treinamento de RNA assim como a utilização de algoritmos genéticos como alternativa.

Esta aplicação se resume em avaliar curvas de carga típicas dos setores residencial, comercial e industrial. A avaliação destas curvas é de fundamental importância à gestão da rede de distribuição de energia elétrica.

No capítulo 6 são apresentadas as conclusões referentes ao trabalho desenvolvido.

Um comentário geral sobre os resultados obtidos assim como sugestões para futuros trabalhos são os assuntos tratados neste capítulo.

O Capítulo 7 traz as referências bibliográficas que serviram de base a realização da presente dissertação.

CAPÍTULO II

2 - REDES NEURAIS ARTIFICIAIS

2.1 INTRODUÇÃO

São modelos computacionais que trabalham de forma similar ao cérebro humano. Assim como a estrutura do sistema nervoso humano, as redes neurais artificiais (RNA) são formadas por elementos de processamento operando paralelamente, chamados de neurônios, interligados por conexões nervosas chamadas sinapses [01].

A este conjunto de neurônios fortemente conectados, pode-se associar algum tipo de conhecimento que será armazenado nas conexões existentes entre os neurônios adjacentes.

Ao processo de associação de conhecimento à RNA dá-se o nome de Aprendizado, que é atribuído à RNA pela execução de um outro processo chamado Treinamento.

Após “treinar” a RNA ela estará apta a realizar as tarefas para as quais ela foi treinada.

Algumas das principais características das RNA são:

- Capacidade de Aprendizagem por Exemplos
- Adaptabilidade
- Capacidade de Generalização
- Tolerância a falhas
- Boa imunidade a ruídos

Algumas das características acima citadas justificam o motivo desta técnica ser amplamente utilizada nas mais diversas aplicações.

A seguir serão dados maiores detalhes sobre projeto, desenvolvimento e implementação de uma RNA.

2.2 HISTÓRICO DAS REDES NEURAIS ARTIFICIAIS

A história das redes neurais artificiais teve seu início com a publicação do primeiro estudo desenvolvido por Warren McCulloch e Walter Pitts. No artigo intitulado *A logical calculus of the ideas immanent in nervous activity*, publicado em 1943, McCulloch e Pitts propuseram um modelo simplificado para a representação do neurônio biológico [07]. Este estudo teve como base o fato de que, num dado instante de tempo, o neurônio pode estar ativo ou inativo de forma a propagar ou não a informação nele contida. O trabalho por eles desenvolvido estimulou o estudo de novos modelos neurais por pesquisadores de grande representatividade.

Alguns anos depois, em 1949 Donald O. Hebb em seu livro *The Organization of Behavior* (a organização do comportamento) definiu o conceito de atualização dos pesos sinápticos [08]. Hebb definiu alguns conceitos fundamentais como o fato de que numa rede neural a informação está contida nos pesos sinápticos e que estes são alterados no processo de aprendizagem da rede neural.

No decorrer de quase uma década os resultados apresentados no estudo de novos modelos neurais não foram de considerada importância até que, em 1958 em livro *Principles of Neurodynamics* (Princípios de Neurodinâmica), Frank Rosenblatt apresenta o primeiro modelo de rede neural chamado de Perceptron [09].

O modelo desenvolvido por Rosenblatt é extremamente simples, sendo composto por uma camada de entrada e uma camada de saída. A cada unidade na camada de entrada está associado um “peso sináptico” ou simplesmente “peso” (denominação que será adotada ao longo desta dissertação). Na camada de saída teremos a soma ponderada da entrada, ou seja, a soma de cada entrada multiplicada pelo seu respectivo peso [02]. Como definido por McCulloch e Pitts o neurônio pode assumir apenas dois estados: ativo e inativo, por isso foi

incorporada ao modelo uma função limiar capaz de estabelecer uma fronteira entre estes dois estados.

O detalhamento deste modelo e dos seus sucessores será apresentado posteriormente.

No início da década seguinte, mais especificamente em 1962, Bernard Widrow e Marcian Hoff desenvolveram o ADALINE (Adaptative Linear Element) e o MADALINE (Many ADALINE) Perceptron.

Estes dois novos modelos capazes de resolver tarefas relacionadas ao reconhecimento de padrões trabalham com entradas analógicas ao invés de binárias como no modelo proposto por McCulloch e Pitts.

No ano de 1969, Marvin Minsky e Seymour Papert do Massachusetts Institute of Technology - MIT, publicaram o famoso livro chamado *Perceptrons* [06], onde os autores analisaram a rede Perceptron e mostraram suas limitações. Uma das principais limitações apontada no livro faz menção ao fato de que a rede Perceptron se aplica apenas a problemas linearmente separáveis e que a mesma seria incapaz de resolver problemas simples como o XOR (função Ou-Exclusivo). Esta publicação causou grande repercussão e teve grande impacto na comunidade científica da época, fazendo com o que o estudo de novas técnicas e modelos fossem condenados ao desinteresse por vários anos.

Embora em descrença, uma contribuição expressiva foi dada em 1972 por Teuvo Kohonen da Universidade de Helsinky que propôs o modelo chamado de *Self-Organizing Feature Map* (Mapa Auto-Organizável de Características), onde os valores de entrada, dos pesos e das saídas podem ser contínuos ao invés de binários. Além disso, em seu modelo, ao invés de se utilizar um único neurônio na saída optou por utilizar vários neurônios para representar o padrão de entrada. Estes neurônios são influenciados pelo que mais se

assemelha ao padrão de entrada, ou seja, não há apenas um neurônio na camada de saída responsável por gerar a saída da rede neural, mas sim um grupo interagindo entre si.

Até 1982 os modelos desenvolvidos tinham a característica de serem *Feedforward*, ou seja, o fluxo do sinal era num único sentido, com o sinal na entrada se propagando em direção à saída. Neste mesmo ano o físico John Hopfield apresenta um novo modelo conhecido como Modelo de Hopfield, onde o sinal de entrada se propaga em direção à saída e em seguida é realimentado à entrada, caracterizando uma rede *Feedback* e fazendo com que redes neurais passem a atuar como sistemas dinâmicos.

Em 1986 Geoffrey Hinton, David Rumelhart e Ronald Williams apresentam o algoritmo de treinamento denominado *Backpropagation Error* [10], baseado na retropropagação do erro gerado na saída em direção às unidades de entrada, para o ajuste dos pesos da RNA. Isto possibilitou a adição de uma camada intermediária à rede Perceptron.

O algoritmo apresentado calcula a diferença entre a saída da rede e um valor desejado pré-estabelecido, gerando assim um sinal de erro. Este sinal de erro é propagado para a camada anterior permitindo o ajuste dos pesos dessa camada, juntamente com uma função não-linear, o que faz com que a rede seja capaz de representar funções não-lineares como a função “ou-exclusivo”.

Isto fez, do algoritmo Backpropagation, associado à rede perceptron com camadas intermediárias, o modelo mais utilizado em todos os tempos pela sua capacidade de representação de uma grande gama de padrões.

A apresentação deste modelo estimulou o interesse de um grande número de pesquisadores pela área de redes neurais e conseqüentemente o desenvolvimento de modelos cada vez mais aperfeiçoados, encurtando assim a distância entre o biológico e o artificial.

2.3 O NEURÔNIO ARTIFICIAL

O neurônio artificial é tido como a principal unidade de uma rede neural artificial. Ele foi totalmente inspirado no neurônio biológico sendo constituído por unidades de entrada, um conjunto de valores chamados de “pesos sinápticos” ou simplesmente “pesos” (que são associados a cada uma das entradas) e unidades de saída, conforme mostrado na figura abaixo:

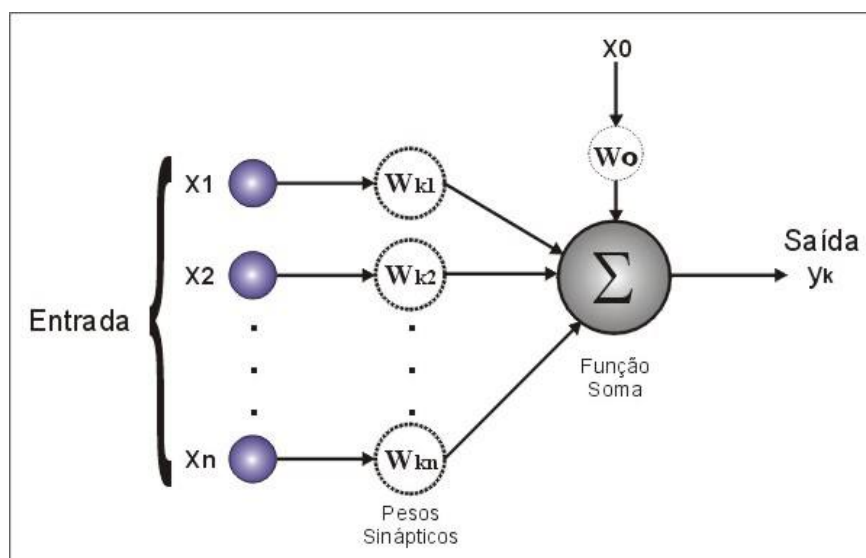


Figura 2.1: Modelo de um neurônio artificial proposto por McCulloch e Pitts

Os valores apresentados às unidades de entrada ($x_1, x_2, x_3, \dots, x_n$) são multiplicados pelos seus respectivos pesos ($w_{k1}, w_{k2}, w_{k3}, \dots, w_{kn}$). Desta forma, é simples notar que o peso expressa a “força de ligação” entre dois neurônios quaisquer [04]. O valor da entrada x_0 é sempre igual a 1 e o peso associado a esta entrada é denominado “bias”. Os valores das entradas multiplicadas pelos respectivos pesos são apresentados à função soma que gera o sinal de saída y_k , expresso da seguinte forma:

$$y_k = w_0 + x_1 * w_{k1} + x_2 * w_{k2} + x_3 * w_{k3} + \dots + x_n * w_{kn}$$

Normalmente os valores dos pesos da rede neural são armazenados em um vetor, e este vetor está associado à memória da rede. Quando ocorre o processo de treinamento eles têm seus valores modificados constantemente e por este fato costuma-se dizer que o “conhecimento” adquirido pela rede está contido nos pesos. Seguindo este raciocínio pode-se dizer que o “aprendizado” se dá quando todos estes pesos têm seus valores ajustados para gerar uma saída pré-determinada.

2.4 PERCEPTRON

No intuito de permitir que pequenas variações na entrada do neurônio possam ser percebidas na saída, Rosenblatt propôs a inserção de uma “função de ativação” capaz de reproduzir na saída estas pequenas alterações na entrada, amplificando ou atenuando os sinais que irão gerar a saída y_k do neurônio artificial proposto por McCulloch e Pitts.

Este modelo foi chamado de Perceptron e a função de ativação a ele incorporada pode ser de vários tipos como apresentado no tópico a seguir.

A figura a seguir ilustra o modelo proposto por Rosenblatt de acordo com [01].

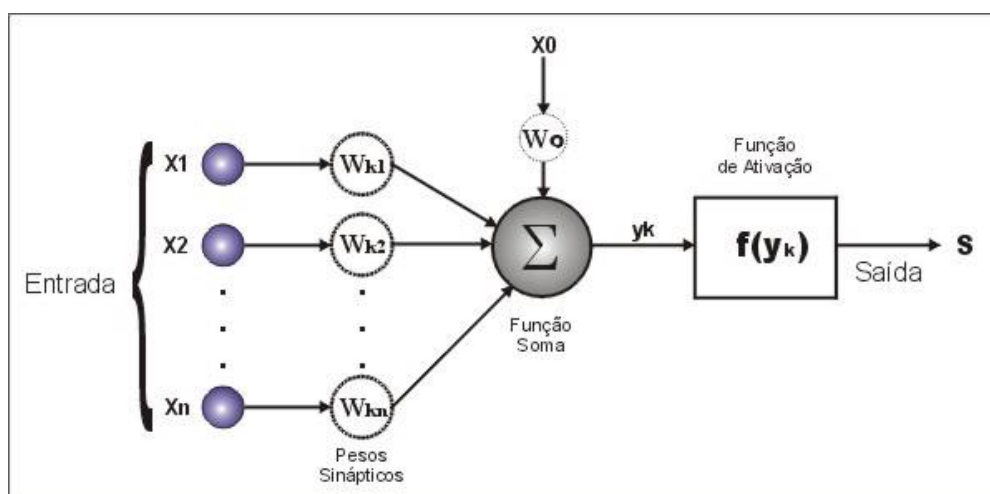


Figura 2.2: Modelo do Perceptron proposto por Rosenblatt

De acordo com a figura 2.2 nota-se que a cada entrada existe um valor associado, e assim como no modelo de McCulloch, este é também chamado de “*peso*”.

Quando um determinado peso assume um valor positivo dizemos que há uma sinapse de natureza excitatória, em caso contrário é comum dizer que ocorre uma sinapse inibitória.

No elemento central, encontra-se a “*função soma*” que realiza o somatório dos resultados de todas as entradas multiplicadas pelos seus respectivos pesos, gerando uma única saída expressa da seguinte forma:

$$y_k = \sum_{i=1}^n (W_i * X_i) + x_0 * w_0$$

A função de ativação tem papel de extrema importância no modelo apresentado. Ela é responsável por limitar a amplitude do sinal na saída do neurônio, sendo normalmente a saída da seguinte forma:

$$S = f(y_k)$$

Nota-se que a saída está condicionada a uma função.

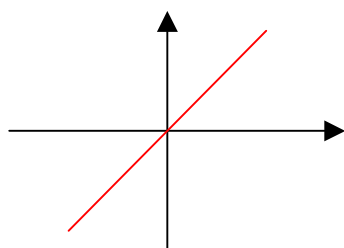
Outra observação importante é que o peso w_0 (*bias*) é somado ao argumento da função de ativação, causando assim um deslocamento da mesma ao longo do seu eixo horizontal.

É também importante ressaltar que o seu valor também deve ser ajustado ao longo do processo de treinamento.

2.5 FUNÇÃO DE ATIVAÇÃO

A função de ativação tem o papel de limitar os sinais de saída num intervalo de valores conhecidos. Vários tipos de funções são utilizados e a escolha da função de ativação adequada é determinada pelo tipo de aplicação. Dentre as mais comuns, são ilustradas abaixo as funções

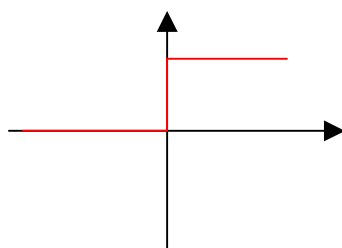
lineares e sigmóides (logarítmicas e hiperbólicas são as mais utilizadas) com suas respectivas características.



Função Linear

Relação Entrada/Saída:

$$a=n$$

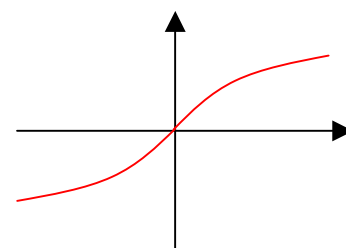


Função Degrau

Relação Entrada/Saída:

$$a=0 \quad n<0$$

$$a=1 \quad n\geq 0$$



Função Tangente Hiperbólica (Sigmóide)

Relação Entrada/Saída:

$$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$$

2.6 TOPOLOGIA

Os neurônios que compõe o sistema nervoso biológico são interligados de diversas maneiras, formando uma complexa rede. Da mesma forma, a RNA pode assumir distintas topologias, algumas mais simples e outras com um certo grau de complexidade.

Em princípio considera-se a rede, na sua forma típica, contendo as entradas, uma camada intermediária chamada de camada oculta e a camada de saída. A topologia mais comum e mais empregada é aquela semelhante a um grafo orientado, com todos os nós conectados a todos os outros. Desta forma é conhecida a rede neural artificial denominada MLP (multi-layer perceptron network), ou rede de perceptrons multi-camadas como ilustrado a seguir:

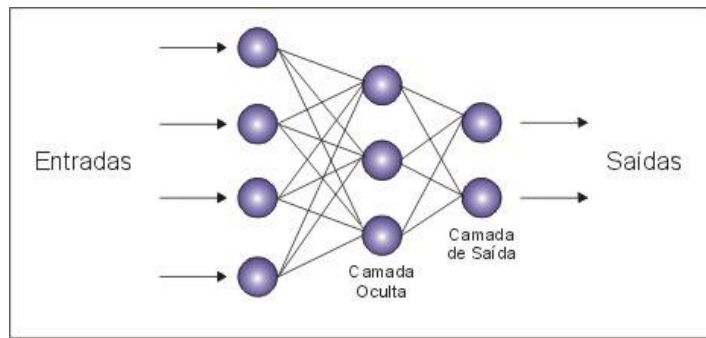


Figura 2.3: Representação Básica de uma rede MLP

Nesta topologia cada neurônio na rede é conectado a todos os outros da camada posterior e/ou anterior. O sinal presente em cada entrada é transferido a todos os neurônios na camada intermediária e assim sucessivamente até a camada de saída.

Nota-se que neste tipo de rede o sinal trafega apenas em um sentido, da entrada para a saída. Redes com esta característica são classificadas como redes diretas (ou “feedforward”). Existem também situações onde o sinal na saída pode ser realimentado para a entrada da rede, constituindo as chamadas redes recorrentes (ou “feedback”). Estes tipos de rede não são o objeto de estudo do presente trabalho não sendo aqui detalhadas.

Uma idéia da topologia de uma rede recorrente é dada pela figura a seguir:

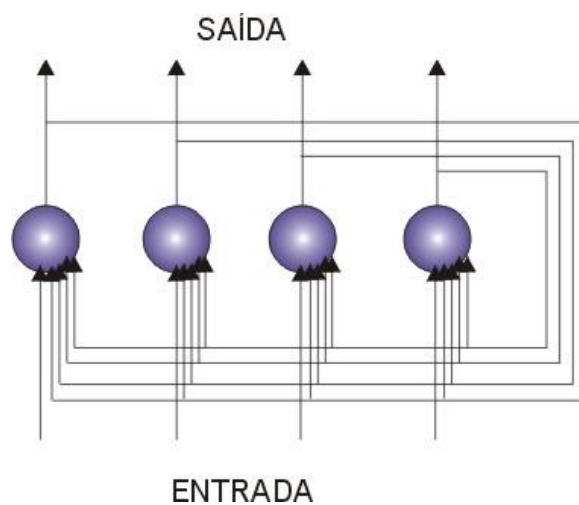


Figura 2.4: Representação Básica de uma rede MLP Recorrente

Nota-se que aos sinais na saída da rede servem também como estímulos para a entrada, este é o grande diferencial das redes recorrentes.

2.7 APRENDIZADO

Aprender é um processo que permite associar um tipo de comportamento de acordo com uma nova situação. Uma das formas de aprendizado está relacionada a situações semelhantes vividas no passado, chamado aprendizado por exemplos.

As RNA têm a capacidade de aprender por exemplos, através de situações vividas anteriormente, assim como realizar “interpolações” entre elas.

No aprendizado conexionista não se procura estabelecer regras de aprendizado e sim determinar a intensidade de conexões entre os neurônios. O processo de aprendizado consiste em determinar o conjunto ótimo de valores que representam de forma correta a intensidade entre as conexões, fazendo com que a rede responda conforme o desejado. A este processo é dado o nome de aprendizado. Pode-se considerar que o conhecimento adquirido está representado neste conjunto pesos entre as conexões, o que faz a RNA estar ligada ao que se chama de conexionismo [05].

De maneira geral o aprendizado pode ocorrer de forma supervisionada e não supervisionada.

2.7.1 Aprendizado Supervisionado

O aprendizado é dito supervisionado quando há conhecimento dos sinais desejados na saída da rede. Quando ocorre o processo de treinamento da RNA pares de entradas e saídas são apresentados a ela. A rede produz uma saída de acordo com cada entrada a ela apresentada.

Se a resposta apresentada for diferente da desejada a RNA gera um sinal de *erro*. Este *erro* pode ser a média do erro quadrático (MSE) ou o somatório do erro quadrático (SSE). No intuito de minimizar este erro, um processo iterativo ajusta os valores dos pesos da rede até que este seja o mínimo possível.

Quando se atinge o estado em que ele assume seu valor mínimo diz-se que a rede está “treinada” e a partir daí será capaz de responder corretamente aos estímulos de entrada.

2.7.2 Aprendizado Não Supervisionado

O aprendizado não-supervisionado não requer o conhecimento prévio das saídas. No processo de treinamento a rede recebe em sua entrada vários tipos de padrões e os organiza em classes arbitrárias. Daí o fato das RNA serem consideradas como classificadoras de padrões.

Quando um padrão é apresentado à entrada, após o treinamento, a saída indica a classe a qual este padrão pertence. Caso o padrão não seja do conhecimento da rede esta cria uma nova classe específica a este tipo de padrão. Este tipo de aprendizado não é muito utilizado sendo mais comum em sistemas em que se deseja apenas a classificação de padrões.

2.8 REDES COM MÚLTIPLAS CAMADAS

As redes com múltiplas camadas (MLP) foram concebidas para resolver problemas que o perceptron não era capaz. Como dito anteriormente, o perceptron se restringe a resolver apenas problemas linearmente separáveis, sendo incapaz de solucionar uma simples operação XOR (ou-exclusivo).

A MLP é o tipo de rede mais utilizada devido à sua baixa complexidade e fácil implementação. São do tipo “*feedforward*”, ou seja, o sinal trafega em sentido único, da entrada para a saída da rede. Podem conter duas camadas intermediárias, sendo muito mais

comuns aplicações com apenas uma camada intermediária. O método de treinamento mais utilizado para este tipo de rede é o “*backpropagation*”, sendo empregadas também variantes deste método que fazem uso de informações do vetor gradiente para a minimização do erro gerado na saída da rede.

O vetor gradiente indica a direção de crescimento de uma determinada função, como ilustrado nos gráficos abaixo:

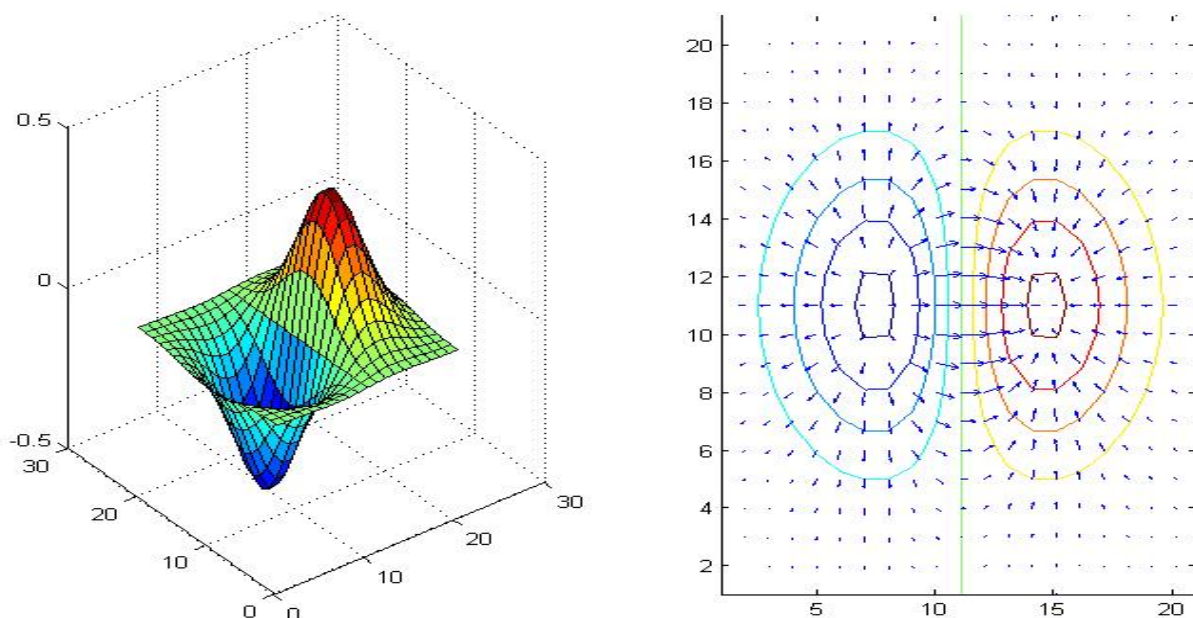


Figura 2.5 : Ilustração de uma superfície qualquer e seus vetores gradientes

O gráfico da esquerda representa uma função qualquer e o da direita é a representação dos vetores gradientes (também chamados de vetores de direção). É importante notar no segundo gráfico que os vetores apontam sempre para a direção de crescimento da função. Desta forma, quando invertemos o sinal do vetor gradiente, este passará a apontar sempre para a direção de decrescimento da função e será chamado de gradiente descendente. Este será utilizado como medida de referência da variação do erro gerado no processo de treinamento das RNA.

No tópico seguinte serão dados maiores detalhes de como ocorre o processo de treinamento através do backpropagation.

2.9 BACKPROPAGATION

Conforme dito anteriormente a rede MLP trabalha com pares de entradas e saídas desejadas para ajustar os pesos da rede. Para o treinamento da rede, o algoritmo *Backpropagation* realiza duas etapas: uma chamada “forward”, e outra “backward”.

Na primeira, um padrão qualquer é apresentado às unidades de entrada, processado nas camadas seguintes gerando uma saída correspondente.

Na segunda fase do treinamento a saída produzida é comparada ao valor desejado e o erro é computado para este padrão. Em seguida os pesos nas unidades de saída são ajustados de acordo com o gradiente descendente do erro em função do conjunto dos pesos da saída. O ajuste dos pesos na camada intermediária e entrada são calculados da mesma maneira, sempre tomando como referência o erro na saída da rede que é retropropagado até a entrada, daí o motivo desta fase ser chamada de “backward”.

Este processo é realizado iterativamente para cada novo padrão apresentado na entrada. Quando todos os padrões tiverem sido apresentados diz-se que uma “época” foi completada. Para se completar o treinamento de uma RNA pode ser necessário um grande número de épocas, isso dependerá do número e da complexidade dos padrões de entrada.

2.9.1 Considerações importantes

Apesar das facilidades de implementação do algoritmo pode-se dizer que nem sempre se obterá sucesso no treinamento da rede. Um dos pontos mais relevantes é que o *Backpropagation* utiliza informações de gradiente para minimização da função erro e em

regiões de gradientes próximos de zero (platôs) ocorre a *paralisia da rede*, ficando o treinamento “estacionado” numa região de mínimo local.

A superfície de erro de uma RNA é complexa, multimodal (com muitos picos e depressões). Quando a rede “estaciona” em um ponto de mínimo local deixa de buscar a depressão mínima mais profunda (mínimo global), que seria o erro mínimo da rede.

Chegando nesta situação entende-se que o erro mínimo foi atingido e o processo de treinamento é interrompido. Isto constitui uma convergência prematura e faz com a rede não responda de maneira correta. Muitos fatores influenciam no sucesso do treinamento de uma RNA, dentre eles os mais importantes são os valores iniciais dos pesos da rede e o conjunto de treinamento. É usual utilizar-se de uma distribuição normal para arbitrar os valores iniciais destes pesos.

CAPÍTULO III

3 - ALGORITMOS GENÉTICOS

3.1 INTRODUÇÃO

Atualmente questões relativas à otimização de processos através da minimização de uma função “custo” ou da maximização de uma função “lucro” têm se tornado cada vez mais comum em várias áreas. Seja em finanças, engenharia e até mesmo medicina, minimizar lucros, custos, erros e uma série de outras variáveis pode significar substanciais aumentos nos ganhos, aumento de produção, de precisão e conseqüentemente na evolução das ferramentas matemáticas associadas aos processos de otimização. Para tipos de problemas como estes foram desenvolvidas técnicas que compõe a chamada Computação Evolucionária, fundamentadas nas teorias que definem a evolução das espécies ao longo das gerações.

A técnica de computação evolucionária adotada neste trabalho é conhecida por algoritmo genético simples, ou simplesmente AG [14][15], proposta por John H. Holland em 1975 no trabalho denominado “Adaptation in Natural and Artificial Systems” [15].

De uma maneira bastante sucinta, os algoritmos genéticos (AG) podem ser entendidos como ferramentas de otimização capazes de minimizar e maximizar funções, baseados na Teoria da Evolução do inglês Charles Darwin que combina genética e seleção natural.

Em outras palavras, os AG são modelos computacionais capazes de processar possíveis soluções para um determinado problema em estruturas semelhantes ao cromossomo biológico e aplicar a estas estruturas operadores genéticos capazes de preservar e também aproximar valores aleatórios em soluções em potencial.

3.2 INSPIRAÇÃO BIOLÓGICA

A base dos AG é o princípio da seleção natural e sobrevivência do mais apto, proposto no ano de 1859 pelo naturalista e fisiologista inglês Charles Darwin em seu livro “A Origem das Espécies”.

A teoria de Darwin propõe que o indivíduo mais bem adaptado ao seu meio ambiente terá mais chances de sobreviver e gerar descendentes.

Apoiados nisto, vários pesquisadores desenvolveram após anos de pesquisa, o Neodarwinismo, que explica como as espécies evoluem ao longo do tempo.

O Neodarwinismo afirma que a história de existência de vida ao longo das gerações, em nosso universo, é regida por alguns poucos processos estatísticos que atuam sobre populações e espécies. Estes processos são definidos como competição, seleção, reprodução e mutação.

Competição e seleção são processos conseqüentes de uma população compartilhando um espaço finito. Por questão da própria sobrevivência estes fatores são impostos diretamente pela natureza.

A reprodução é um processo natural inerente a qualquer ser vivo, e dividida em sexuada e assexuada. A primeira se dá pela união de dois indivíduos diferentes (denominados macho e fêmea), onde o material genético cedido por cada um é organizado em pares de unidades mais simples (cromossomo). Já na reprodução assexuada não há diferença entre genética entre os seres de mesma espécie no que se refere à reprodução. Ocorre, na reprodução sexuada, o chamado “crossover” que é a troca de parte do material genético dos progenitores, gerando um novo material genético, fruto de uma recombinação dos anteriores.

O processo de mutação é a resultante das condições impostas pelo meio, e tem grande importância dentro do processo evolutivo, pois garante a diversidade da população.

3.3 O ALGORITMO GENÉTICO

O algoritmo genético é um processo aleatório através do qual pode-se maximizar (ou minimizar) uma determinada função. Em termos práticos, otimizar consiste em encontrar uma solução que corresponda a um ponto de máximo (ou de mínimo) de uma determinada função. Esta é chamada de função objetivo, ou *função aptidão* segundo a terminologia dos AG.

Ao conjunto de possíveis soluções para a *função aptidão* é dado o nome de *população*. A cada possível solução é dado o nome de *indivíduo* (ou *cromossomo*, segundo a terminologia dos AG). Cada cromossomo é composto por um conjunto de valores, denominados *genes* e o valor que cada gene pode assumir é chamado *alelo*.

O processo inicia-se com a criação de população inicial de forma aleatória. Esta população inicial é composta por número de indivíduos pré-determinado. Estes indivíduos nada mais são do que conjuntos de números gerados aleatoriamente considerados como soluções em potencial para o problema em questão. O passo seguinte é avaliar a “aptidão” de cada um destes indivíduos em relação à função aptidão. A cada indivíduo avaliado é atribuído um valor de aptidão de acordo com sua “proximidade” da solução do problema. Em outras palavras, quanto mais próximo da solução, maior a aptidão atribuída a este indivíduo e, conseqüentemente, maior a chance de suas características (genótipos) serem “herdadas” por seus descendentes nas futuras gerações.

Após a avaliação de todos os indivíduos, serão escolhidos aqueles com maiores aptidões para compor a nova população. Esta nova população é gerada através da aplicação de *operadores genéticos* nos indivíduos da população atual. Quando esta nova população estiver constituída diz-se que foi completada uma geração.

Esta nova população é avaliada, seus melhores indivíduos irão constituir uma próxima geração e assim por diante.

Este processo é repetido iterativamente até que a solução seja encontrada ou então o número de gerações especificado seja atingido. O fluxograma a seguir ilustra o processo:

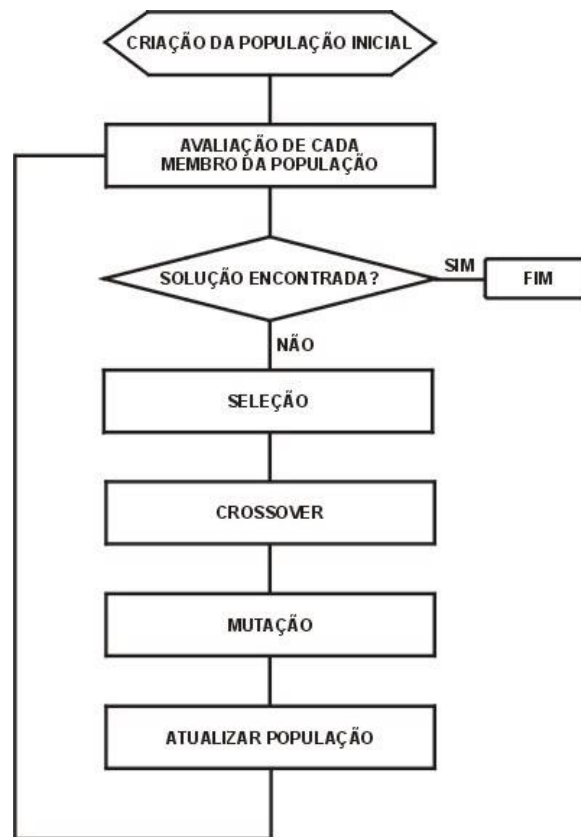


Figura 3.1: Fluxograma Básico dos Algoritmos Genéticos [02]

3.3.1 População Inicial

Criar uma população inicial significa definir um conjunto de valores (indivíduos), de uma forma pré-definida, que serão candidatos à solução do problema em questão.

A forma de criação da população inicial é de grande importância para o sucesso do processo. Para isto é fundamental o conhecimento do espaço de busca ou região factível, que normalmente corresponde ao domínio do problema em questão.

Conhecendo-se estes limites máximos e mínimos que determinam o espaço de busca é possível então gerar, segundo uma distribuição normal, uma quantidade de valores aleatórios igual ao número de indivíduos escolhido.

Outro ponto fundamental é definir como serão representados os valores gerados aleatoriamente. Na maioria dos casos cada um destes valores é representado por uma *string binária*, correspondente ao valor original. Quando se utiliza a notação binária, pode-se dizer que o alfabeto utilizado para a codificação dos indivíduos foi o binário, enquanto que se optarmos por representar os indivíduos por números reais, é correto dizer que foi utilizado o alfabeto real. Como exemplo adotaremos um indivíduo com valor igual a 0,637197 e outro com valor igual a 10, conforme a tabela a seguir:

<i>Indivíduo</i>	<i>Valor</i>	<i>Codificação Binária:</i>	<i>Codificação Real:</i>
1	0,637197	'1000101110110101000111'	0,637197
2	10	'0000000000000000001010'	10

Tabela 3.1: Exemplos de Codificações

Para o exemplo anterior optou-se por uma codificação binária com comprimento igual a vinte e dois bits. O valor binário correspondente ao primeiro indivíduo corresponde ao número decimal 2288967 mapeado nos limites do espaço de busca escolhido [-1;2]. Este mapeamento foi feito de acordo com a seguinte expressão:

$$\text{Indivíduo} = [\text{min} + (\text{max} - \text{min}) * b_{10}] / (2^L - 1)$$

O valor L corresponde ao comprimento do número binário codificado e b_{10} ao valor na base 10. Os valores “min” e “max” correspondem aos extremos do intervalo que determina o espaço de busca.

A escolha do Alfabeto para a codificação da população inicial depende muito da aplicação da técnica, visto que na maioria dos casos é usado o alfabeto binário, porém o alfabeto real vem sendo cada vez mais utilizado por facilitar a implementação além de reduzir drasticamente o custo computacional. Maiores detalhes sobre codificação serão dados posteriormente neste trabalho.

O passo seguinte consiste em avaliar, de acordo com a função aptidão, cada indivíduo da população inicial e atribuir um valor de aptidão compatível com sua proximidade da solução ideal.

3.3.2 Função Aptidão

A função aptidão é de grande importância para o sucesso do algoritmo pois permite avaliar indivíduos candidatos a solução e também determinar sua permanência na população que irá compor a geração seguinte.

Na maioria dos casos definir a função aptidão é uma tarefa simples, uma vez que ela será a própria função que se deseja otimizar.

Após definir a função aptidão deve-se avaliar cada indivíduo da população e então definir a possibilidade de cada um deles permanecer nas gerações futuras através de um parâmetro chamado probabilidade de seleção, ou simplesmente *PS*.

A probabilidade de seleção indica qual a chance de um indivíduo, em relação a todos os outros, de compor a próxima geração.

Matematicamente, a probabilidade de seleção pode ser definida como:

$$PS_i = \frac{f_i}{\sum_{i=1}^n f_i}$$

Depois de calculada a probabilidade de seleção de cada indivíduo, utiliza-se o método de seleção conhecido por “roda da roleta” para selecionar os indivíduos que irão compor a população intermediária. Este processo de seleção simula o papel da seleção natural na evolução, selecionando para sobreviver e reproduzir os indivíduos mais bem adaptados ao meio, no caso do AG aqueles com maiores valores para a função aptidão.

A população intermediária é criada para alocar os indivíduos selecionados pela roda da roleta, denominados pais e que serão submetidos aos operadores genéticos para gerar a nova população.

No método de seleção os indivíduos são dispostos como em uma roleta, de acordo com seu respectivo valor de probabilidade de seleção. A figura a seguir ilustra esta situação.

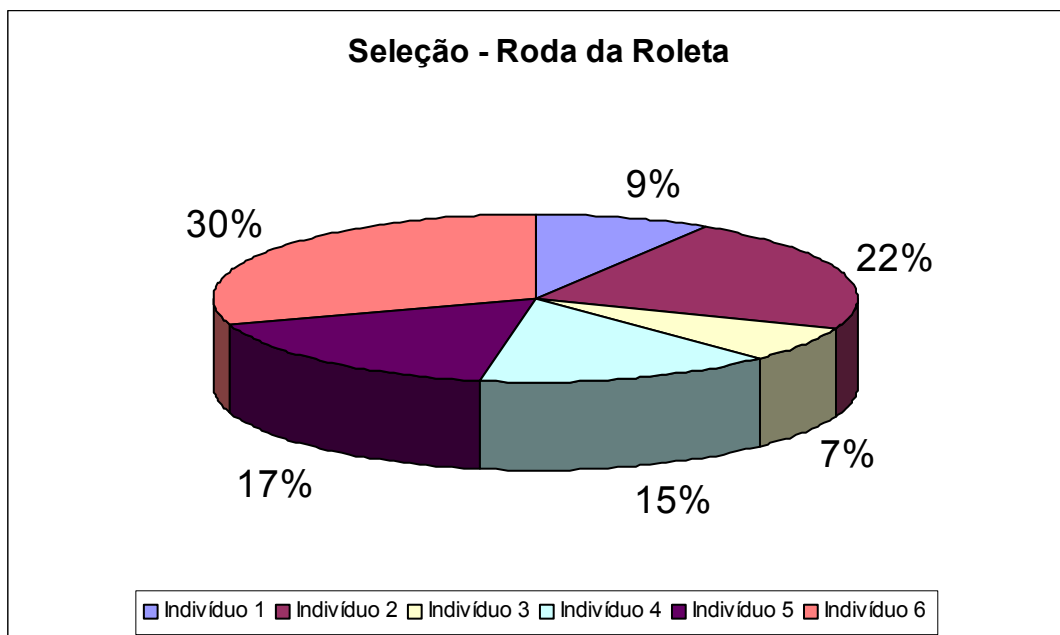


Figura 3.2 – Seleção pelo Método da Roleta

O processo de seleção consiste em gerar um número aleatório entre zero e o somatório das probabilidades de seleção de todos os indivíduos, de acordo com uma distribuição normal, simulando o movimento real de uma roleta.

Este número aleatório estará contido dentro do intervalo pertencente a um dos indivíduos, sendo este o escolhido pelo método. Este processo é repetido até que o número de indivíduos na população intermediária seja suficiente para a criação da próxima geração que terá sempre um tamanho fixo pré-determinado.

Após o término da composição da população intermediária são aplicados os operadores genéticos, descritos a seguir.

3.3.3 Operadores Genéticos

Conforme dito anteriormente os AG simulam os processos de evolução dos seres vivos, e de acordo com as definições da Teoria da Evolução os indivíduos competem entre si para conseguir alimentos, se reproduzir e até mesmo pela sobrevivência. Nesta condição sabe-se que apenas os mais “fortes” e mais bem adaptados ao ambiente irão sobreviver, reproduzir e gerar descendentes que herdarão suas características, presentes em seu material genético. Quando dois indivíduos de mesma espécie se reproduzem existe a troca do material genético entre eles, o que garantirá a presença de características dos pais em seus descendentes. Este processo, dentro do AG, é chamado de crossover e garante que esta troca seja efetuada ao longo das gerações no processo de evolução.

Outro importante fenômeno presente no processo de reprodução é o aparecimento de características nos filhos que não estão presentes nos pais. Este fenômeno é chamado de mutação e pode se dizer que também é o responsável pela diversidade da população a cada processo de reprodução.

3.3.3.1 Crossover

A troca de material genético entre dois indivíduos se dá pelo processo conhecido por *crossover*. Neste processo o indivíduo gerado pelo cruzamento de dois outros é formado por parte do material genético de ambos. Esta troca proporciona, na prática, a exploração pontos totalmente novos no espaço de busca. Este tipo de exploração é denominado *exploration*.

A figura a seguir ilustra a forma como acontece o tipo mais comum conhecido por Crossover de Um Ponto.

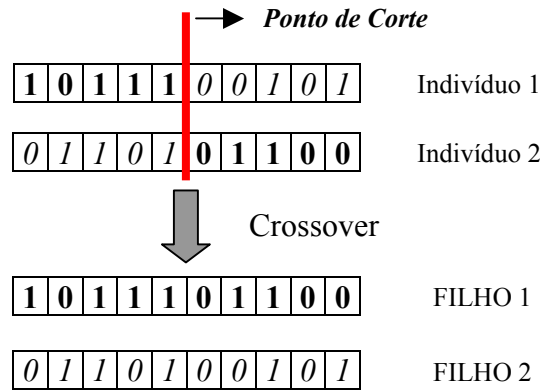


Figura 3.3: Exemplo de crossover para a codificação binária

Ponto de Corte é o local, escolhido aleatoriamente, que limita a porção do material cedido por cada indivíduo.

Para a Codificação Real deve-se optar por um tipo de crossover próprio para este tipo de codificação, uma vez que os métodos convencionais podem não adicionar nenhum novo tipo de informação.

Na literatura são encontradas várias opções de crossover para codificação real, dentre eles os mais utilizados são o crossover média e o blend crossover, também conhecido por BLX- α . Cada um destes tipos é responsável por explorar o espaço de busca de diferentes formas. A seguir é dado um exemplo de crossover média.

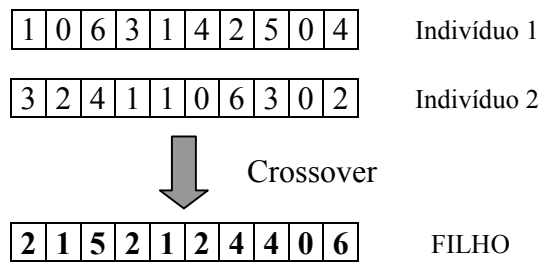


Figura 3.4: Exemplo de crossover para a codificação real

Neste tipo de crossover o filho é resultado da média aritmética entre os genes dos cromossomos pais. Isto faz com que haja uma forte tendência dos genes serem levados para o centro do espaço de busca, causando assim uma perda de diversidade. Por este motivo deve-se optar por outro tipo, como o blend crossover, por exemplo. Esta escolha é feita em função do tipo de problema em questão.

3.3.3.2 Mutação

O processo conhecido por mutação consiste em acrescentar aos indivíduos gerados (filhos) características diferentes daquelas presentes em seus “pais”. Isto garante um processo chamado *exploitation* que, na prática, é responsável por explorar o espaço de busca utilizando informações de pontos anteriormente visitados a fim de encontrar melhores pontos.

A mutação acontece através da inserção de alguma nova característica ao indivíduo gerado, o que pode ser feito dentro do AG invertendo-se o valor de um bit (gene) por exemplo. No caso da codificação real basta trocar o valor de um gene por outro gerado aleatoriamente.

Codificação Binária

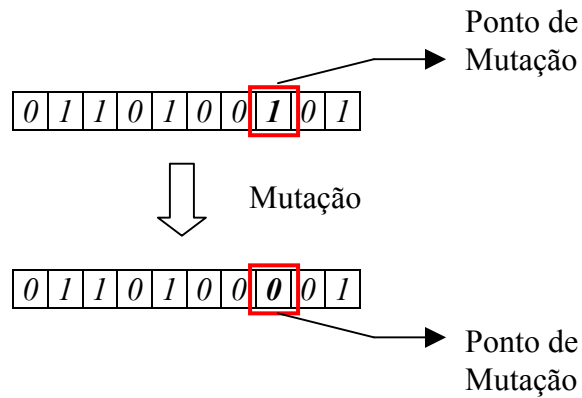


Figura 3.5: Exemplo de mutaçao para a codificaçao binária

Codificaçao Real

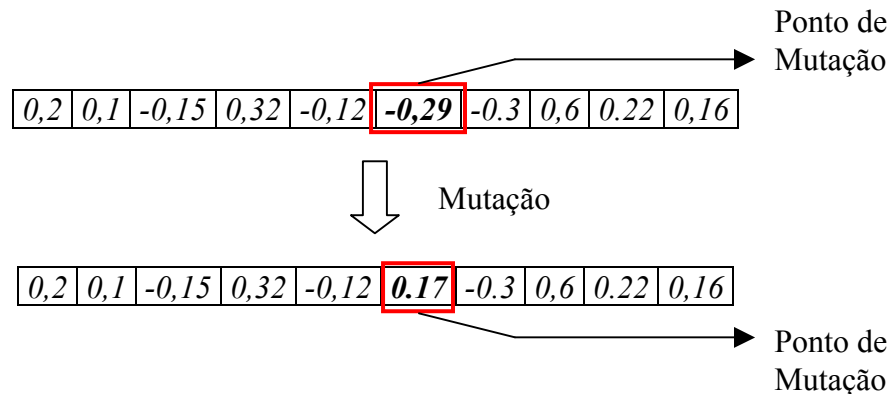


Figura 3.6: Exemplo de mutaçao para a codificaçao real

O Ponto de Mutaçao é onde ocorre a modificaçao do gene e é escolhido aleatoriamente.

3.3.4 Critérios de Parada

Após a composiçao da populaçao intermediária e aplicaçao dos operadores genéticos têm-se os novos indivíduos (filhos) gerados que irão compor a nova populaçao.

Esta nova populaçao passa pelos mesmos processos de avaliaçao, seleçao e reproduçao.

À cada criaçao de uma nova populaçao diz-se que passou uma *geraçao*. Este processo é

repetido iterativamente até que a solução (indivíduo ótimo) seja encontrada ou então que o número de gerações pré-estabelecido foi atingido. Normalmente são estes os critérios que determinam o fim da execução do AG.

3.3.5 Parâmetros Importantes

Para a execução do AG alguns parâmetros fundamentais devem ser especificados. Os mais importantes e aqui apresentados são o número de indivíduos que irão compor a população inicial, a taxa de crossover, a taxa de mutação e o número máximo de gerações através do qual o algoritmo irá “percorrer”.

Estes parâmetros podem ser ajustados durante a execução do algoritmo, porém estes ajustes em tempo de execução podem comprometer o resultado final, desta forma optou-se por ajustá-los inicialmente e mantê-los fixos durante o processo.

A seguir serão explicados estes parâmetros assim como as modificações causadas por alterações em seus valores.

3.3.5.1 Número de Indivíduos

O número de indivíduos é um dos principais parâmetros a ser ajustado e tem grande influência na execução do algoritmo.

A escolha do número de indivíduos é bastante particular, ou seja, em problemas onde o número de variáveis a serem otimizadas for pequeno não há grandes variações em termos de tempo de processamento, porém nos casos em que se deseja otimizar um grande número de variáveis, elevações no número de indivíduos podem implicar em aumentos consideráveis no tempo de processamento.

Desta forma a escolha do número de indivíduos deve ser feita experimentalmente, levando-se em conta que um número muito pequeno pode conduzir a uma convergência

prematura assim como um número muito grande pode aumentar consideravelmente o tempo de execução do algoritmo.

3.3.5.2 Taxa de Crossover

O crossover é o mecanismo que permite a busca da melhor solução através da combinação de porções de cromossomos pais. No entanto este processo deve acontecer de forma controlada, fazendo com que não se perca alguma informação importante contida em alguns indivíduos com altas aptidões. Diversos valores são sugeridos na literatura para a taxa de crossover, porém os mais comuns estão no intervalo entre 0.6 e 0.65.

3.3.5.3 Taxa de Mutação

Conforme dito anteriormente, o processo de mutação é responsável por adicionar ao cromossomo características novas, de forma que todas as regiões do espaço de busca possam ser exploradas. Porém este deve ser também um processo que ocorre de forma controlada, para que não haja perda de material genético e comprometa a aptidão do cromossomo em questão. Assim como acontece com a taxa de crossover, diversos valores são sugeridos na literatura para a taxa de mutação, estando os mais comuns entre 0.001 e 0.05.

CAPÍTULO IV

4 - SISTEMAS HÍBRIDOS

4.1 INTRODUÇÃO

Com a consolidação das técnicas de Inteligência Artificial e suas constantes aplicações nas mais diversas áreas, uma nova geração de sistemas inteligentes começou a ser desenvolvida através da agregação duas ou mais técnicas de inteligência artificial. Quando há este tipo de associação de técnicas, dizemos que o sistema desenvolvido é híbrido.

A hibridização de técnicas pode resultar em poderosos sistemas onde uma técnica pode suprir a deficiência da outra, aproveitando-se apenas das vantagens que cada uma delas oferece em separado.

Diversos exemplos de hibridização de técnicas podem ser encontrados em [3], como exemplos sistemas desenvolvidos com a agregação de sistemas especialistas com lógica fuzzy, algoritmos genéticos e sistemas fuzzy, redes neurais e sistemas fuzzy, etc.

4.2 SISTEMA HÍBRIDO PROPOSTO

4.2.1 Considerações Iniciais

Conforme dito no tópico anterior, o motivo de se associar técnicas é aproveitar as vantagens de cada uma e desenvolver sistemas cada vez mais precisos e consistentes.

Neste trabalho em específico a proposta é o treinamento de redes neurais artificiais (RNA) através de algoritmos genéticos para o reconhecimento de padrões contidos em curvas de consumo de energia elétrica.

Quando ocorre o treinamento de uma RNA os pesos sinápticos são ajustados para que a RNA apresente uma saída pré-determinada. Tal tarefa não é tão simples assim e vários fatores podem fazer com que a RNA não responda de maneira desejada.

Muitos motivos podem fazer com que isto aconteça, por exemplo, a super especialização e o não reconhecimento de padrões, causados pela inconsistência do conjunto de treinamento. Outra questão importantíssima e já citada é a definição do número de neurônios na camada intermediária.

O ponto chave no treinamento de um RNA é o comportamento do erro de saída da RNA pois é em função dele que os pesos sinápticos são ajustados. À medida que ocorre o treinamento os pesos são ajustados de forma a minimizar o erro. Este ajuste é feito, normalmente, por algoritmos de treinamento que utilizam o vetor gradiente.

O vetor gradiente indica a direção de crescimento de uma função contínua qualquer, porém se invertemos seu sinal, passará a “apontar” para a direção de decrescimento da função em questão e será chamado de gradiente descendente.

A utilização do gradiente descendente para treinamento de um RNA pode não produzir valores satisfatórios quando a superfície de erro for multimodal e o treinamento estacionar num ponto de mínimo local, deixando de atingir o mínimo erro na saída da rede.

A sugestão para a solução deste problema é a utilização de algoritmos genéticos que também são aplicados à minimização de funções.

A vantagem do uso de AG é justamente o fato de não utilizarem nenhum tipo de informação fornecida pelo gradiente, não havendo assim a possibilidade do treinamento estacionar em algum ponto de mínimo local.

Desta forma propõe-se realizar os treinamentos necessários utilizando o Algoritmo de Retropropagação do Erro (Backpropagation), assim como os AG para que se possa traçar um paralelo comparativo entre os métodos aplicados ao reconhecimento de padrões.

4.2.2 O Sistema Híbrido

Para que se possa implementar um sistema híbrido deve-se conhecer as particularidades de cada técnica em separado assim como o comportamento assumido quando ambas operam em conjunto.

No caso do treinamento de RNA através de AG, pontos como topologia da rede, número de neurônios, erro máximo permitido, tamanho da população do AG, critérios de parada, alfabeto de codificação, tipos de crossover e mutação, assim como uma série de outros parâmetros devem ser escolhidos de maneira a otimizar os resultados finais.

Como foram realizados inúmeros treinamentos todos os parâmetros específicos de cada treinamento serão apresentados em separado no capítulo seguinte.

Após a determinação de todos os parâmetros necessários, será dado início ao treinamento que se desenvolverá conforme a seqüência executada pelo AG.

Abaixo são ilustrados os dois casos em que ocorre o treinamento pelos dois métodos: gradiente descendente e algoritmo genético.

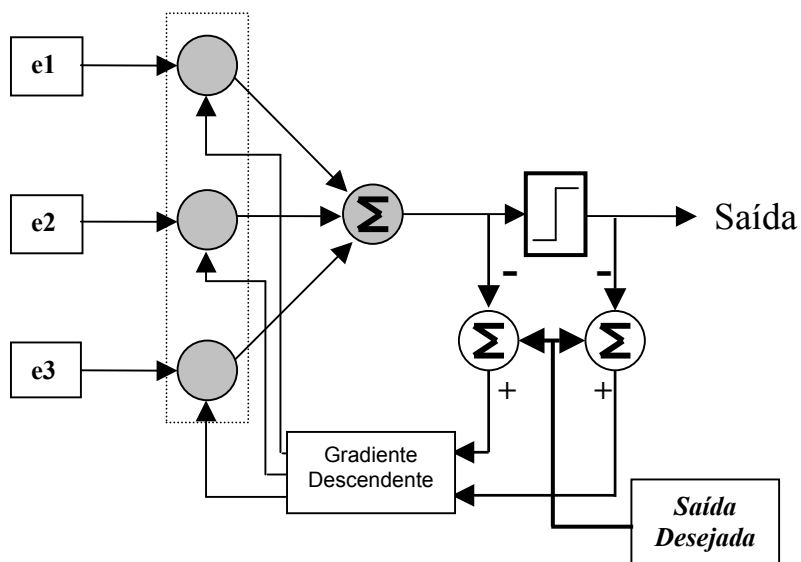


Figura 4.1 - Treinamento da RNA através de métodos baseados em Gradiente Descendente

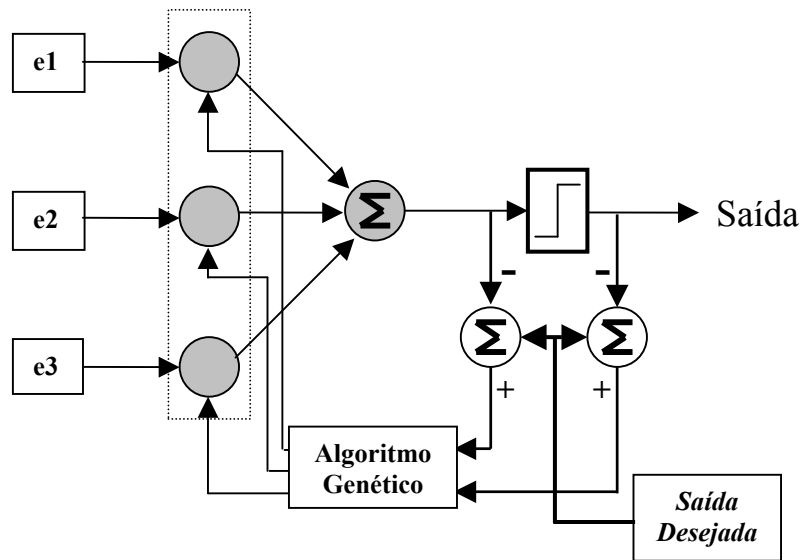


Figura 4.2 - Treinamento da RNA através de Algoritmos Genéticos

4.2.3 Etapas do Processo

Para que se possa iniciar o treinamento da RNA, deve primeiramente determinar sua topologia, o número de neurônios em cada uma das camadas, a função de ativação dos neurônios e obviamente os conjuntos de treinamento e validação.

Definidas as características principais da RNA parte-se para as definições dos parâmetros do AG.

O número de genes que formarão os cromossomos é exatamente igual ao número de ligações entre os neurônios mais o número de bias de todos os neurônios da RNA.

Suponha, por exemplo, que se deseja treinar uma RNA com vinte neurônios na camada de entrada, dez na camada intermediária e dois na camada de saída, e que cada padrão de treinamento é representado por um vetor coluna com vinte e cinco linhas.

Neste caso pode-se dizer que o cromossomo terá o seguinte comprimento:

$$L = (25*20) + (20*10) + (10*2) + 20 + 10 + 2 = 752 \text{ genes}$$

Onde L representa o comprimento do cromossomo.

Pode-se dizer que cada cromossomo será composto por setecentos e cinquenta e dois genes e que o algoritmo genético interpretará cada um destes genes como uma variável. Analisando desta forma pode-se dizer que se trata de um problema de otimização onde o se deve encontrar um valor ideal para cada uma destas variáveis de forma que a RNA apresente, ao final do treinamento, o menor erro possível e que no processo de validação seja capaz de reconhecer padrões até então desconhecidos.

Após conhecer o comprimento de cada cromossomo deve-se gerar a população inicial através de um número de indivíduos pré-determinado. Conforme dito anteriormente estes indivíduos serão gerados de acordo uma distribuição normal com limites inferiores de superiores também pré-definidos.

Gerada a população, o próximo passo será avaliar cada indivíduo e classificá-los de acordo com o erro gerado por cada um deles.

Como cada indivíduo corresponde a um conjunto de pesos sinápticos da RNA, basta apresentar o conjunto de treinamento à RNA utilizando estes pesos e calcular o erro em sua saída. Este processo se repete para cada conjunto de pesos, ou indivíduo, e o erro é calculado para cada indivíduo. De posse do erro gerado por todos os indivíduos, passa-se ao processo de Seleção executado pelo AG, onde aqueles que produziram o menor erro serão selecionados e submetidos aos operadores genéticos.

Os operadores genéticos serão responsáveis pela recombinação destes indivíduos assim como a inserção de alguma nova característica afim de que haja exploração de todo o espaço de busca.

Terminada a recombinação dos melhores indivíduos é formada a nova população e o processo se repete até que um indivíduo ótimo seja encontrado ou o número de gerações máximo estipulado seja atingido.

4.3 VANTAGENS DO HIBRIDISMO

Associar técnicas no intuito de obter resultados cada vez melhores pode não ser uma tarefa tão simples assim. À medida que acontece essa associação deve-se considerar um aumento na complexidade do processo como um todo. Cada técnica operando isoladamente possui, inerentemente, uma série de parâmetros a serem ajustados. Quando se decide associá-las, a quantidade de ajustes a serem feitos é muito maior e, o ajuste em uma das técnicas produz mudanças nos resultados produzidos pela outra. Considerando esta situação é óbvio que o hibridismo tenha que oferecer grandes vantagens que justifiquem seu uso.

Apostando nisto, acredita-se que ainda seja conveniente trabalhar com técnicas associadas. Dentre as vantagens proporcionadas, podemos destacar:

- A existência de mínimos locais na superfície de erro não influencia no resultado final, ou seja, não há risco de convergência prematura.
- A complexidade dos padrões de treinamento não afeta o AG, o que não é garantido acontecer com o backpropagation devido aos cálculos de derivadas direcionais realizadas sobre a superfície de erro.

4.4 RESULTADOS ESPERADOS

Quando se propõe o treinamento de RNA através de AG deve-se ter em mente que teremos a partir disto, um problema de otimização com um grande número de variáveis, igual o tamanho do cromossomo calculado anteriormente. Este ajuste de pesos, quando feito pelo backpropagation, é fortemente dependente da superfície de erro. Quando esta for uma superfície multimodal há a possibilidade do algoritmo encontrar um valor de erro mínimo local e interpretá-lo como ótimo, ignorando o valor do erro mínimo global. São consideradas multimodais curvas que apresentam vários pontos de mínimo e/ou de máximo, como ilustrado na figura seguinte:

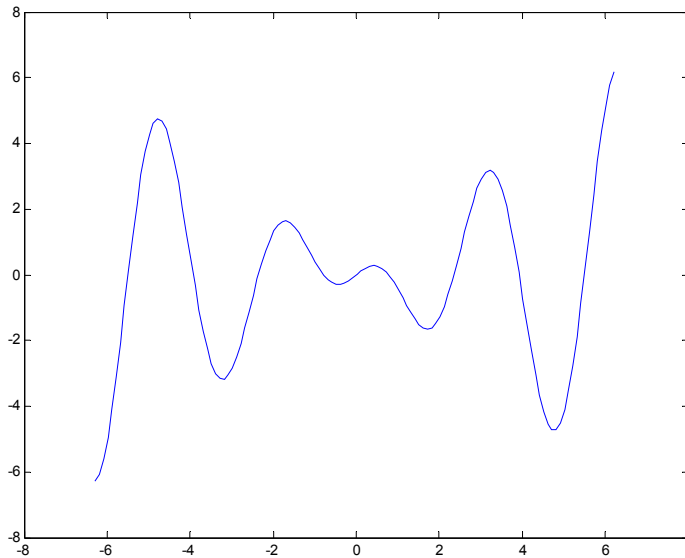


Figura 4.3 - Exemplo de uma Curva Multimodal

Como o AG é um processo aleatório e não utiliza informações de gradiente, espera-se que os resultados por ele apresentados sejam aqueles que realmente resultem no menor valor de erro possível. Não se sabe, porém, se isto procede tanto em problemas de pouca quanto de muita complexidade. A partir da afirmação anterior é que serão testados vários casos de forma que seja possível, ao final deste trabalho, estabelecer um paralelo comparativo entre AG e backpropagation.

CAPÍTULO V

5 - ESTUDO DE CASO

5.1 INTRODUÇÃO

Até o momento foram apresentados os principais conceitos e definições sobre RNA, AG e sistemas híbridos, porém nenhum tipo de aplicação envolvendo estes conceitos foi desenvolvido.

Desde o início deste estudo houve uma grande preocupação com aplicação prática, de forma que este trabalho não se resumisse a um simples comparativo entre técnicas, mas a base para o desenvolvimento de uma poderosa ferramenta computacional para a avaliação do sistema de distribuição de energia elétrica através de informações fornecidas por curvas de duração de carga, ou simplesmente curvas de carga.

As curvas de carga são perfis de demanda em períodos de tempo específicos. De posse da curva de carga de um ponto qualquer do sistema de distribuição de energia elétrica, poder-se-á determinar as porções de consumo correspondentes a cada tipo de setor (residencial, comercial e industrial) presentes na curva fornecida. Isto proporciona uma série de vantagens que serão comentadas posteriormente.

A partir deste ponto serão apresentadas as simulações realizadas que servirão de base para o comparativo final, assim como as conclusões e a contribuição que será disponibilizada aos futuros trabalhos envolvendo sistemas híbridos.

5.2 APLICAÇÃO PRÁTICA

Atualmente existe uma grande preocupação por parte das empresas de energia em oferecer um serviço de qualidade permitindo-lhes aplicação de tarifas compensadoras. Isto só

é possível diante de uma constante supervisão, da disponibilização de recursos operacionais assim como investimentos na rede elétrica de distribuição.

Os atuais modelos de supervisão dos sistemas de distribuição baseiam-se no cálculo de demanda dos transformadores, que é realizado através de uma curva de correlação consumo x demanda (KWh x KVA) . Esta curva é normalmente estimada através de um método numérico chamado de regressão linear simples, o qual pode apresentar pouca precisão em relação à necessária para supervisão da rede elétrica de distribuição [19].

Visando uma melhoria na distribuição, a proposta deste trabalho é desenvolver uma metodologia baseada em Redes Neurais Artificiais (RNA) para o gerenciamento da rede elétrica de distribuição através da análise de curvas de carga características dos tipos mais importantes de consumidores como o setor residencial, comercial e industrial.

Segundo o Atlas da Energia no Brasil [13], estes são os setores mais representativos em termos de consumo de energia elétrica até o ano de 1998, com uma parcela equivalente a 83,8% de toda a energia consumida no país. O gráfico a seguir ilustra esta situação:

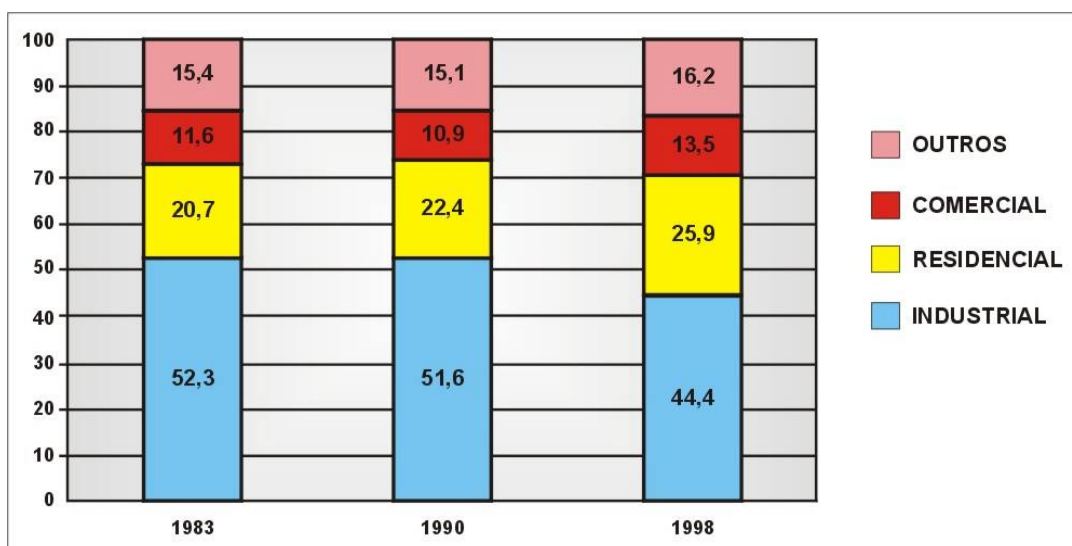


Figura 5.1 - Participação dos Principais Setores de Atividade no consumo de energia elétrica no país (%)

Este cenário veio se modificando até os dias de hoje, porém os setores que ocupam maior parcela de consumo continuam sendo os mesmos e em função disto foram escolhidos para o presente estudo.

A aplicação prática aqui proposta consiste em avaliar curvas de carga de um ponto qualquer da rede de distribuição e identificar a quantidade de energia consumida referente a cada um destes setores.

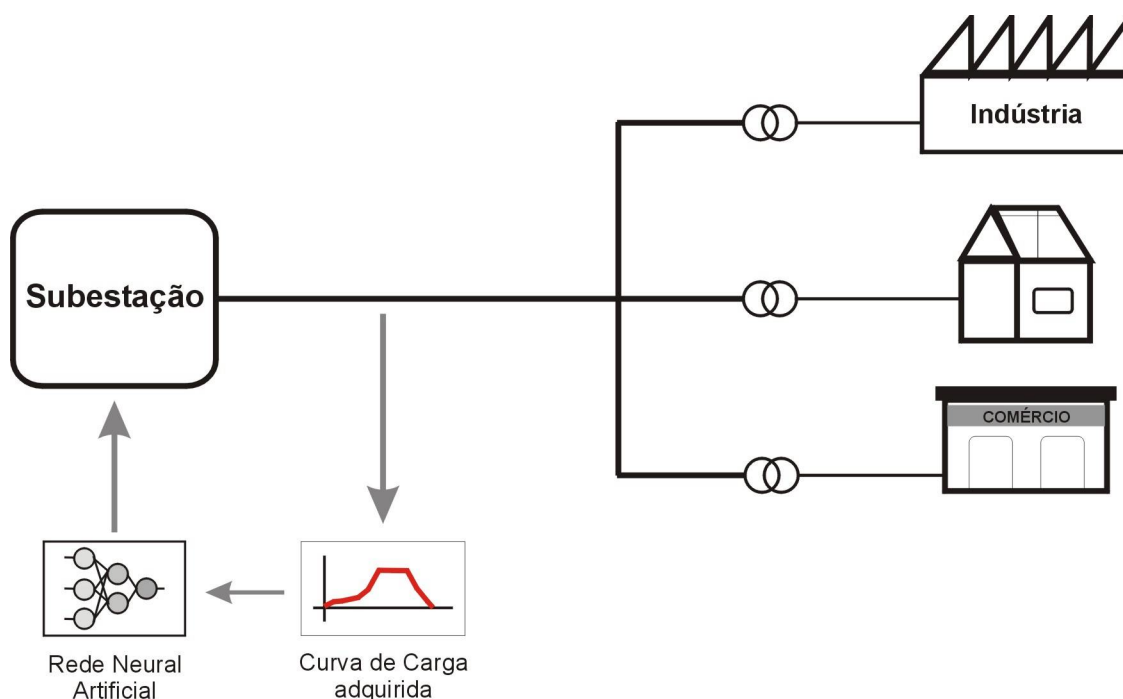


Figura 5.2 – Aplicação Prática

O conhecimento da quantidade de energia consumida por determinado setor permite a alocação da parcela não utilizada a outros setores, ou seja, a redistribuição inteligente pode significar maior disponibilidade.

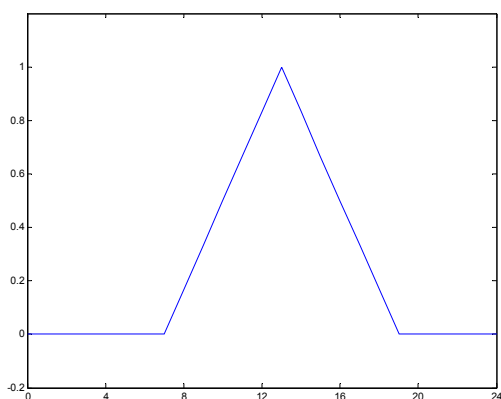
A partir das vantagens acima citadas são iniciadas as simulações necessárias à comprovação da viabilidade desta aplicação. Os testes, simulações realizadas e resultados obtidos estão detalhados nos tópicos seguintes.

5.3 DEFINIÇÃO DOS PADRÕES DE TREINAMENTO

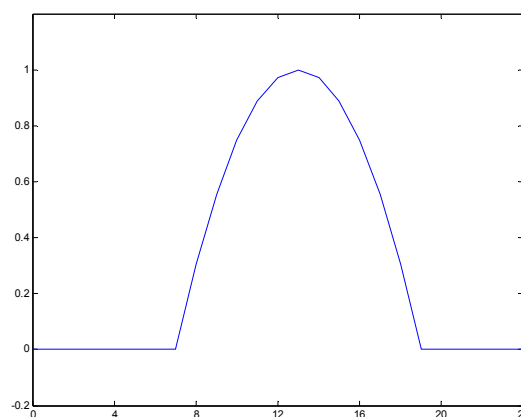
Quando se deseja realizar o treinamento de uma RNA é necessário um conjunto de treinamento capaz de expressar com exatidão os padrões ou conjuntos de padrões que a rede deverá identificar e classificar. Mais importante ainda será garantir a maior diversidade de informações possíveis no conjunto de treinamento, de forma a se evitar que a RNA fique especializada em determinados padrões, não reconhecendo variações destes.

Como o foco do trabalho está na avaliação de padrões contidos em curvas, optou-se primeiramente por observar quais os tipos de resultados apresentaria uma RNA treinada para reconhecer padrões naturais assim como padrões resultantes de associações entre os mesmos.

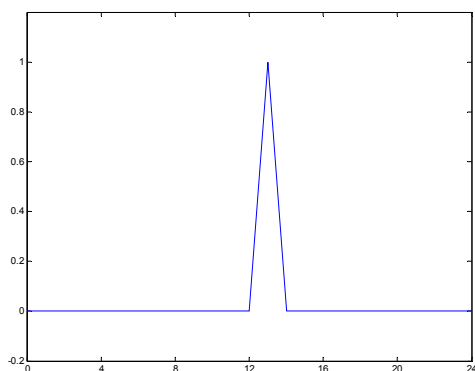
Os padrões naturais escolhidos para a realização dos primeiros testes foram uma onda triangular, uma parábola, um pulso retangular e um impulso, conforme mostrado nos gráficos abaixo:



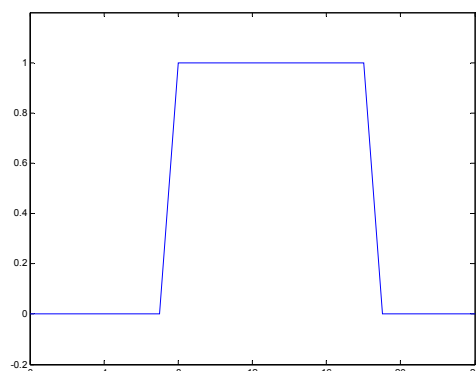
Onda Triangular



Parábola



Impulso



Pulso Retangular

Figura 5.3: *Padrões Naturais*

Nota-se que as escalas das abscissas, determinada pelo intervalo $[0,24]$, foi escolhida com o propósito de representar as vinte e quatro horas diárias. Nas próximas simulações esta mesma escala será utilizada, porém para a representação de curvas de consumo de energia elétrica.

Outro ponto importante é que não há defasagem entre os instantes de tempo em que ocorrem os picos da curva, sendo considerado como ponto principal os valores máximos atingidos.

5.4 SIMULAÇÕES ENVOLVENDO PADRÕES NATURAIS SIMPLES E COMPOSTOS

Definidos os padrões de treinamento, podem-se agora determinar os pontos referentes a RNA em si. Conforme citado anteriormente, vários parâmetros devem ser ajustados de forma que a RNA seja capaz de reconhecer as classes de padrões referentes àqueles para os quais ela foi treinada.

Primeiramente o treinamento é realizado com um conjunto formado apenas por padrões simples e depois é verificado se a RNA é capaz, de reconhecer composições feitas a partir destes.

Desta forma, foram criados conjuntos de treinamentos contendo dois tipos de padrões, com amplitudes variadas, no intuito de verificar se esta RNA, após treinada, reconheceria composições destes padrões com amplitudes diferentes daquelas utilizadas para o treinamento. O primeiro par de padrões escolhido é formado pelo pulso retangular e impulso.

Todas os treinamentos e simulações realizados neste trabalho foram executados com auxílio do software MATLAB, versão 6.5, desenvolvido pela The MathWorks, Inc.

Foram utilizadas também as toolboxes de Redes Neurais Artificiais (Neural Network Toolbox) e de Algoritmos Genéticos (Genetic Algorithm and Direct Search Toolbox) que acompanham o MATLAB. Nesta primeira fase de testes foi utilizada apenas a Toolbox de Redes Neurais como forma de avaliação do conjunto de treinamento e verificação de resultados.

Ao final das simulações são dadas as informações relativas ao treinamento e testes.

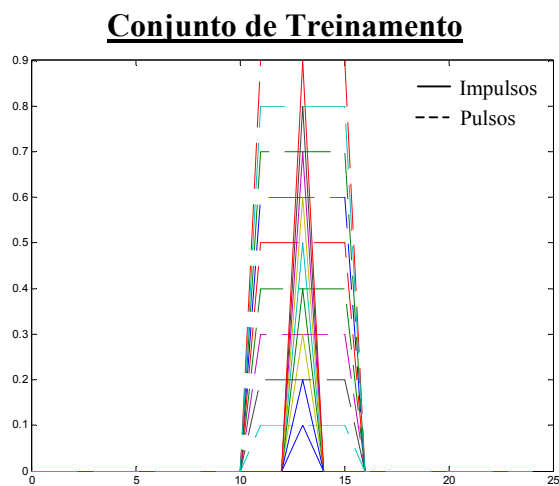


Figura 5.4: Conjunto de treinamento composto por impulsos e pulsos retangulares

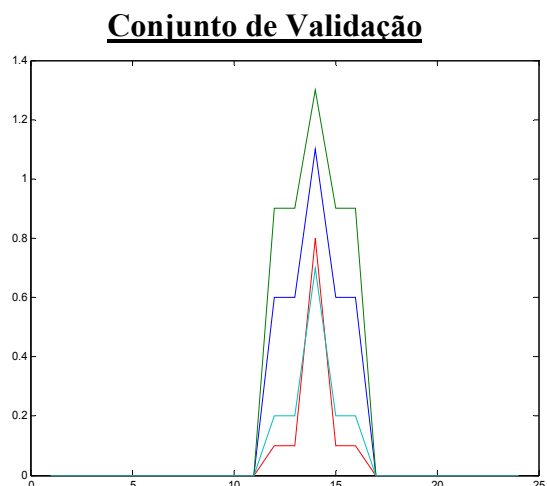


Figura 5.5: Conjunto de validação formado pela composição de impulsos e pulsos retangulares

<i>Padrões Utilizados:</i>																							
<u><i>Treinamento</i></u>	<u><i>Validação</i></u>																						
<p>Impulsos com Amplitudes iguais a:</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>0,1</td><td>0,2</td><td>0,3</td><td>0,4</td><td>0,5</td><td>0,6</td><td>0,7</td><td>0,8</td><td>0,9</td> </tr> </table> <p>Pulsos Retangulares com Amplitudes iguais a:</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>0,1</td><td>0,2</td><td>0,3</td><td>0,4</td><td>0,5</td><td>0,6</td><td>0,7</td><td>0,8</td><td>0,9</td> </tr> </table>	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	<table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>0,5 (I) + 0,6 (R)</td> </tr> <tr> <td>0,4 (I) + 0,9 (R)</td> </tr> <tr> <td>0,7 (I) + 0,1 (R)</td> </tr> <tr> <td>0,5 (I) + 0,2 (R)</td> </tr> </table> <p style="text-align: center;"><u>Onde:</u></p> <p>(R) – representa o Pulso Retangular (I) – representa o Impulso</p> <p><i>Ex: 0,5 (R) indica um Pulso Retangular com amplitude igual a 0,5</i></p>	0,5 (I) + 0,6 (R)	0,4 (I) + 0,9 (R)	0,7 (I) + 0,1 (R)	0,5 (I) + 0,2 (R)
0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9															
0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9															
0,5 (I) + 0,6 (R)																							
0,4 (I) + 0,9 (R)																							
0,7 (I) + 0,1 (R)																							
0,5 (I) + 0,2 (R)																							

Para o conjunto de validação, conforme citado anteriormente, foram utilizadas composições dos padrões de treinamento. Primeiro testou-se a RNA com uma composição de um pulso retangular com amplitude igual a 0.5, somado a um impulso com amplitude igual a 0.6. Depois foi testado um pulso retangular com amplitude igual a 0.4, somado a um impulso com amplitude igual a 0.7 e assim sucessivamente conforme mostrado acima.

A tabela a seguir mostra os resultados esperados nas saídas da RNA assim como aqueles produzidos pela simulação executada.

Resultados Esperados		Resultados Obtidos	
<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>
<i>Impulso</i>	<i>Pulso Retangular</i>	<i>Impulso</i>	<i>Pulso Retangular</i>
0,5	0,6	0,4029	0,8082
0,4	0,9	0,4048	0,9596
0,7	0,1	0,7980	0,1299
0,5	0,2	0,4947	0,2624

Tabela 5.1: Resultados da simulação para padrões como impulsos e pulsos retangulares

Nota-se, após este primeiro treinamento com padrões naturais simples, que a RNA reconhecer padrões compostos, porém com uma certa dificuldade na aproximação do valor de amplitude.

No segundo teste realizado foram escolhidos como padrões de entrada uma parábola e um impulso. Assim como feito na primeira simulação, o conjunto de treinamento criado é formado por diversas amostras destes padrões com variadas amplitudes.

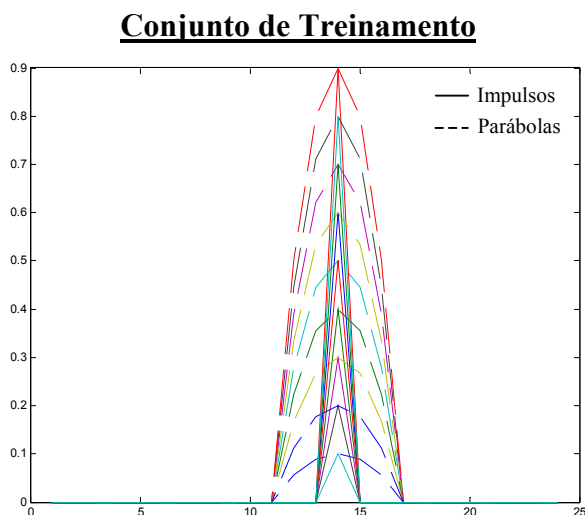


Figura 5.6: Conjunto de treinamento formado por parábolas e impulsos

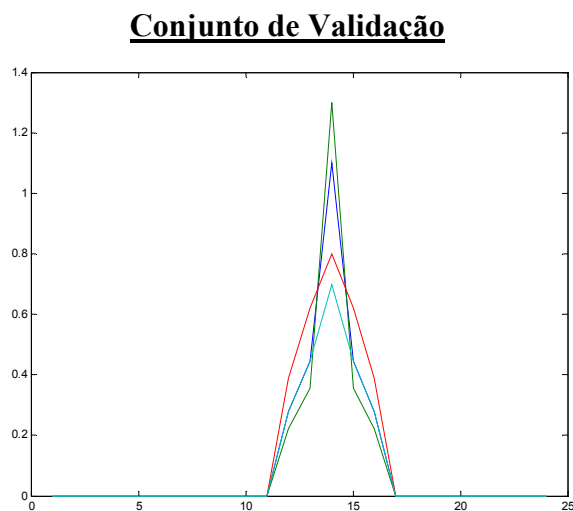


Figura 5.7: Conjunto de validação formado pela composição de parábolas e impulsos

Padrões Utilizados:																							
<p><u>Treinamento</u></p> <p>Parábolas com Amplitudes iguais a:</p> <table border="1" style="margin: auto; border-collapse: collapse;"> <tr> <td>0,1</td><td>0,2</td><td>0,3</td><td>0,4</td><td>0,5</td><td>0,6</td><td>0,7</td><td>0,8</td><td>0,9</td> </tr> </table> <p>Impulsos com Amplitudes iguais a:</p> <table border="1" style="margin: auto; border-collapse: collapse;"> <tr> <td>0,1</td><td>0,2</td><td>0,3</td><td>0,4</td><td>0,5</td><td>0,6</td><td>0,7</td><td>0,8</td><td>0,9</td> </tr> </table>	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	<p><u>Validação</u></p> <table border="1" style="margin: auto; border-collapse: collapse;"> <tr> <td>0,5 (P) + 0,6 (I)</td> </tr> <tr> <td>0,4 (P) + 0,9 (I)</td> </tr> <tr> <td>0,7 (P) + 0,1 (I)</td> </tr> <tr> <td>0,5 (P) + 0,2 (I)</td> </tr> </table> <p>Onde:</p> <p>(P) – representa a Parábola (I) – representa o Impulso</p>	0,5 (P) + 0,6 (I)	0,4 (P) + 0,9 (I)	0,7 (P) + 0,1 (I)	0,5 (P) + 0,2 (I)
0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9															
0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9															
0,5 (P) + 0,6 (I)																							
0,4 (P) + 0,9 (I)																							
0,7 (P) + 0,1 (I)																							
0,5 (P) + 0,2 (I)																							

Resultados Esperados		Resultados Obtidos	
<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>
<i>Parábola</i>	<i>Impulso</i>	<i>Parábola</i>	<i>Impulso</i>
0,5	0,6	0,7416	0,9850
0,4	0,9	0,4268	0,9997
0,7	0,1	0,7279	0,3059
0,5	0,2	0,6915	0,5868

Tabela 5.2: Resultados da simulação para padrões como parábolas e impulsos

Nota-se que, neste caso, os resultados apresentados são muito piores que os gerados pela simulação anterior.

Uma questão relevante quando se trata um impulso como um padrão de treinamento é a possibilidade da RNA interpretá-lo como um ruído. Como uma das características das RNA é a boa imunidade a ruídos, imaginou-se que o padrão pudesse passar despercebido por ela, mas isto não ocorreu.

Para as próximas simulações optou-se por um par de padrões que não envolvessem o Impulso, e neste terceiro caso foram adotados a parábola e a onda triangular. Novamente os valores amplitudes utilizados são os mesmos das simulações anteriores.

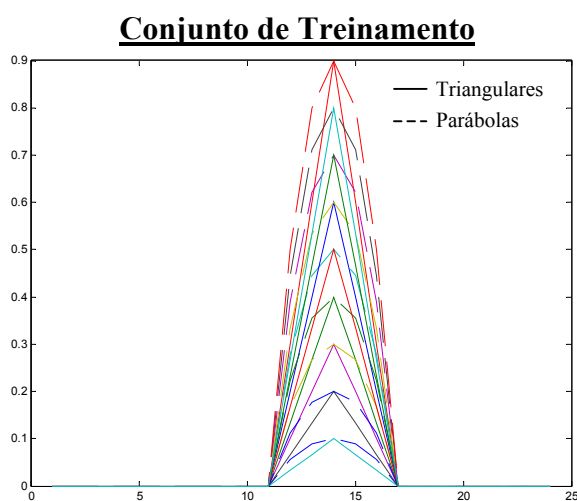


Figura 5.8: Conjunto de treinamento formado pela composição de parábolas e ondas triangulares

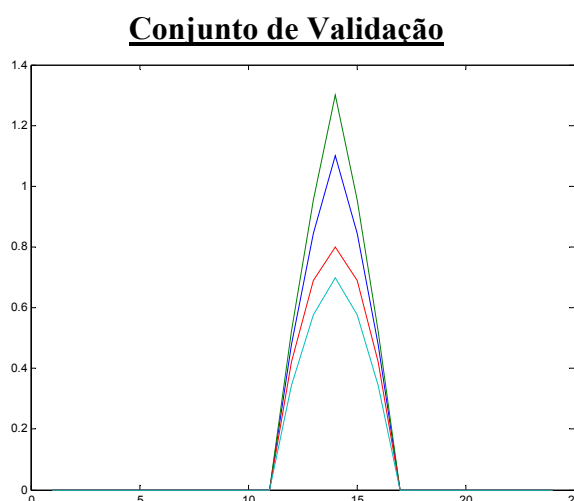


Figura 5.9: Conjunto de validação formado pela composição de parábolas e ondas triangulares

Padrões Utilizados:																							
<u>Treinamento</u>	<u>Validação</u>																						
Parábolas com Amplitudes iguais a: <table border="1" style="width: 100%; text-align: center;"> <tr> <td>0,1</td><td>0,2</td><td>0,3</td><td>0,4</td><td>0,5</td><td>0,6</td><td>0,7</td><td>0,8</td><td>0,9</td> </tr> </table> Ondas triangulares com Amplitudes iguais a: <table border="1" style="width: 100%; text-align: center;"> <tr> <td>0,1</td><td>0,2</td><td>0,3</td><td>0,4</td><td>0,5</td><td>0,6</td><td>0,7</td><td>0,8</td><td>0,9</td> </tr> </table>	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	<table border="1" style="width: 100%; text-align: center;"> <tr> <td>0,5 (P) + 0,6 (T)</td> </tr> <tr> <td>0,4 (P) + 0,9 (T)</td> </tr> <tr> <td>0,7 (P) + 0,1 (T)</td> </tr> <tr> <td>0,5 (P) + 0,2 (T)</td> </tr> </table> Onde: (T) – representa a Onda Triangular (P) – representa a Parábola	0,5 (P) + 0,6 (T)	0,4 (P) + 0,9 (T)	0,7 (P) + 0,1 (T)	0,5 (P) + 0,2 (T)
0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9															
0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9															
0,5 (P) + 0,6 (T)																							
0,4 (P) + 0,9 (T)																							
0,7 (P) + 0,1 (T)																							
0,5 (P) + 0,2 (T)																							

Resultados Esperados		Resultados Obtidos	
<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>
<i>Parábola</i>	<i>Onda Triangular</i>	<i>Parábola</i>	<i>Onda Triangular</i>
0,5	0,6	0,7155	0,7068
0,4	0,9	0,6540	0,8618
0,7	0,1	0,7573	0,0750
0,5	0,2	0,5747	0,1589

Tabela 5.3: Resultados da simulação para padrões como parábolas e ondas triangulares

Os resultados obtidos nesta simulação também não são satisfatórios. Percebe-se que mesmo quando não se utiliza a função impulso como padrão a RNA permanece encontrando dificuldades em reconhecer padrões compostos.

A próxima simulação foi realizada utilizando como padrões de treinamento a onda triangular e o pulso retangular.

Conjunto de Treinamento

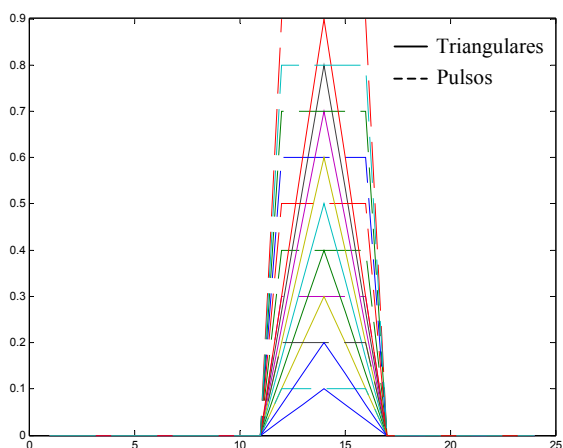


Figura 5.10 Conjunto de treinamento formado por ondas triangulares e pulsos retangulares

Conjunto de Validação

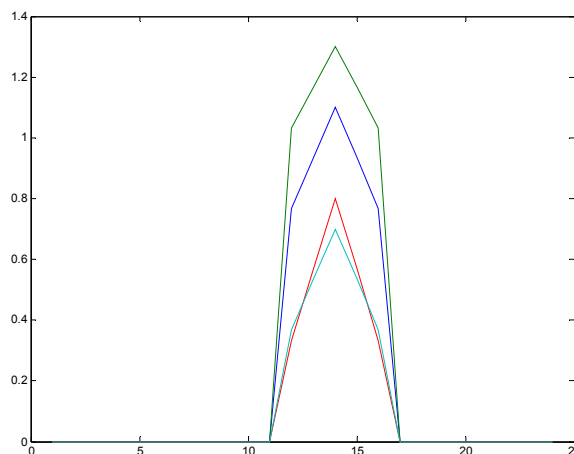


Figura 5.11 Conjunto de validação formado pela composição de ondas triangulares e pulsos retangulares

<i>Padrões Utilizados:</i>	
<u><i>Treinamento</i></u>	<u><i>Validação</i></u>
Ondas Triangulares com Amplitudes iguais a:	
0,1 0,2 0,3 0,4 0,5 0,6 0,7 0,8 0,9	0,5 (T) + 0,6 (R)
	0,4 (T) + 0,9 (R)
	0,7 (T) + 0,1 (R)
	0,5 (T) + 0,2 (R)
Pulsos Retangulares com Amplitudes iguais a:	
0,1 0,2 0,3 0,4 0,5 0,6 0,7 0,8 0,9	<u>Onde:</u>
	(T) – representa o Pulso Triangular
	(R) – representa o Pulso Retangular

Resultados Esperados		Resultados Obtidos	
<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>
<i>Onda Triangular</i>	<i>Pulso Retangular</i>	<i>Onda Triangular</i>	<i>Pulso Retangular</i>
0,5	0,6	0,6966	0,8011
0,4	0,9	0,7322	0,9627
0,7	0,1	0,7467	0,1881
0,5	0,2	0,4904	0,4123

Tabela 5.4: Resultados da simulação para padrões como ondas triangulares e pulsos retangulares

Os resultados obtidos nas simulações anteriores comprovam que a RNA, quando treinada apenas por padrões simples não é capaz de reconhecer padrões compostos.

Inúmeras simulações foram realizadas com diferentes topologias, vários conjuntos iniciais de pesos e com grande número de épocas. Os melhores resultados obtidos estão expostos nas tabelas anteriores.

Desta forma, optou-se por realizar mudanças no conjunto de treinamento, no intuito de obter melhores resultados. O próximo tópico trata esta questão e mostra a expressiva melhoria nos resultados quando o treinamento é realizado com um conjunto de padrões diferentes do utilizado até este ponto.

Os parâmetros ajustados para as simulações estão detalhados na tabela a seguir:

CONFIGURAÇÕES DAS RNA UTILIZADAS NAS SIMULAÇÕES ANTERIORES	
<i>Tipo:</i>	Feedforward
<i>Função de Transferência:</i>	Tangente Hiperbólica
<i>Função de Adaptação:</i>	Gradiente Descendente
<i>Função de Performance:</i>	Somatório do Erro Quadrático - SSE
<i>Topologia:</i>	14:8:2
<i>Número de Entradas:</i>	24
<i>Número de Épocas:</i>	300

Tabela 5.5: Parâmetros das RNA

5.4.1 Mudanças no Conjunto de Treinamento

Após a realização das simulações anteriores pode-se notar que as mudanças nas configurações dos parâmetros da RNA não surtiram efeitos necessários à geração de resultados satisfatórios. Desta forma optou-se pela formação de novos conjuntos de treinamento, de forma que estes incorporem algum tipo de informação referente à composição de padrões.

A iniciativa de associar alguns padrões compostos ao conjunto de treinamento tem o objetivo de “ensinar” a rede estas composições, assim como identificar novos padrões conjugados até então desconhecidos por ela.

A partir deste raciocínio os novos conjuntos de treinamento serão formados por dezoito amostras de padrões simples, com variações de amplitude, e oito de padrões compostos.

O conjunto de validação será formado por quatro amostras de padrões conjugados, diferentes daqueles utilizados para o treinamento.

Realizadas as modificações espera-se que a RNA possa ser capaz de identificar composições de padrões, independente do tipo de treinamento utilizado. Em todas as

simulações seguintes serão avaliados os resultados produzidos pelas RNA treinadas por AG e backpropagation como forma de comparar o desempenho de ambas.

Assim como feito anteriormente, foram escolhidos dois tipos de padrões com amplitudes variadas para compor o conjunto de treinamento, porém agora sendo adicionadas composições destes.

A primeira simulação foi realizada utilizando-se pulsos retangulares e impulsos como padrões. Abaixo são dadas informações referentes ao treinamento assim como os resultados obtidos.

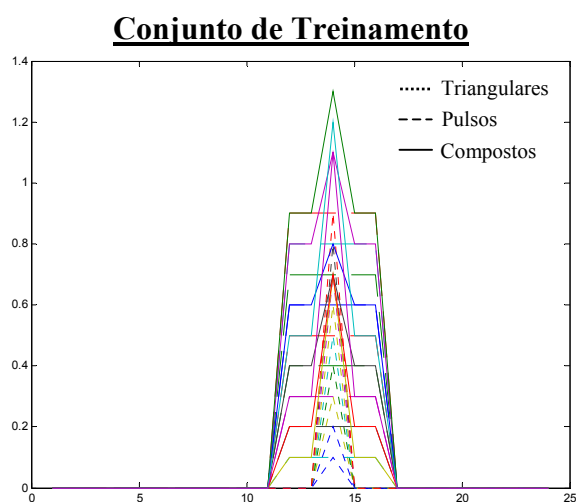


Figura 5.12: Conjunto de treinamento formado por impulsos e pulsos retangulares, assim como composições destes

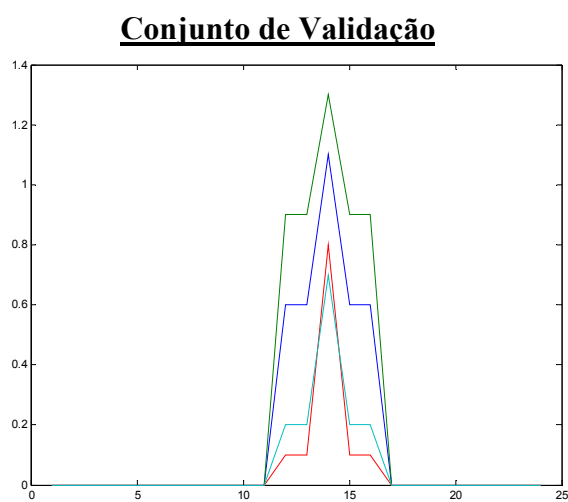


Figura 5.13: Conjunto de validação formado por composições de impulsos e pulsos retangulares

Padrões Utilizados:	
<u>Treinamento</u>	<u>Validação</u>
Impulsos com Amplitudes iguais a:	
0,1 0,2 0,3 0,4 0,5 0,6 0,7 0,8 0,9	
Pulsos Retangulares com Amplitudes iguais a:	
0,1 0,2 0,3 0,4 0,5 0,6 0,7 0,8 0,9	
Somas de Pulsos Retangulares e Impulsos com Amplitudes iguais a:	
0,3 (I) + 0,8 (R) 0,6 (I) + 0,1 (R)	0,5 (I) + 0,6 (R)
0,3 (I) + 0,4 (R) 0,2 (I) + 0,6 (R)	0,4 (I) + 0,9 (R)
0,4 (I) + 0,9 (R) 0,5 (I) + 0,2 (R)	0,7 (I) + 0,1 (R)
0,7 (I) + 0,5 (R) 0,8 (I) + 0,3 (R)	0,5 (I) + 0,2 (R)

Resultados Esperados		Resultados Obtidos			
		BACKPROPAGATION		AG	
<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>
<i>Impulso</i>	<i>Pulso Retangular</i>	<i>Impulso</i>	<i>Pulso Retangular</i>	<i>Impulso</i>	<i>Pulso Retangular</i>
0,5	0,6	0,4963	0,6162	0,6008	0,5760
0,4	0,9	0,4000	0,9000	0,3865	0,7034
0,7	0,1	0,7088	0,0998	0,7118	0,1213
0,5	0,2	0,5000	0,2000	0,5688	0,1373

Tabela 5.6: Resultados da simulação para padrões como impulsos e pulsos retangulares

Nota-se, nesta primeira simulação realizada, que houve uma grande melhoria nos resultados. Tanto a RNA treinada pelo backpropagation quanto pelo AG apresentaram valores satisfatórios em suas saídas. O desempenho do backpropagation, tanto em tempo de execução quanto em resultado final, foi melhor que do AG.

As próximas simulações utilizarão parábolas e impulsos, e assim como foi feito anteriormente, serão adicionadas amostras de composições destes padrões no conjunto de treinamento.

Conjunto de Treinamento

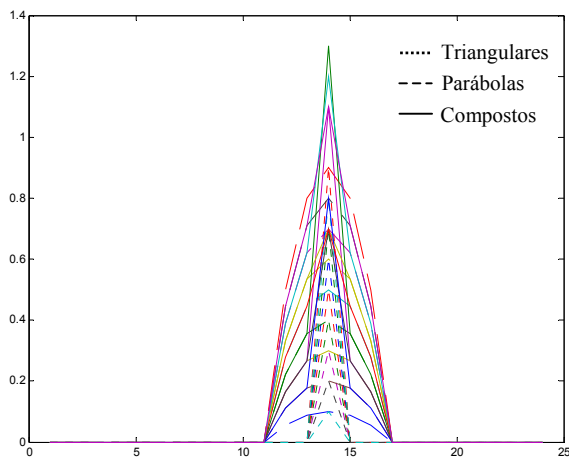


Figura 5.14: Conjunto de treinamento formado por parábolas e impulsos, assim como composições destes

Conjunto de Validação

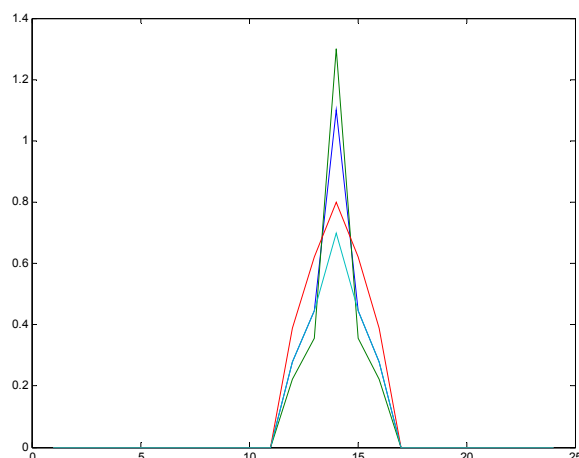


Figura 5.15: Conjunto de validação formado por composições de parábolas e impulsos

Padrões Utilizados:														
<u>Treinamento</u>	<u>Validação</u>													
Parábolas com Amplitudes iguais a: <table border="1" style="margin: auto;"> <tr> <td>0,1</td><td>0,2</td><td>0,3</td><td>0,4</td><td>0,5</td><td>0,6</td><td>0,7</td><td>0,8</td><td>0,9</td> </tr> </table>	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	<table border="1" style="margin: auto;"> <tr><td>0,5 (P) + 0,6 (I)</td></tr> <tr><td>0,4 (P) + 0,9 (I)</td></tr> <tr><td>0,7 (P) + 0,1 (I)</td></tr> <tr><td>0,5 (P) + 0,2 (I)</td></tr> </table>	0,5 (P) + 0,6 (I)	0,4 (P) + 0,9 (I)	0,7 (P) + 0,1 (I)	0,5 (P) + 0,2 (I)
0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9						
0,5 (P) + 0,6 (I)														
0,4 (P) + 0,9 (I)														
0,7 (P) + 0,1 (I)														
0,5 (P) + 0,2 (I)														
Impulsos com Amplitudes iguais a: <table border="1" style="margin: auto;"> <tr> <td>0,1</td><td>0,2</td><td>0,3</td><td>0,4</td><td>0,5</td><td>0,6</td><td>0,7</td><td>0,8</td><td>0,9</td> </tr> </table>	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9					
0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9						
Somas de Parábolas e Impulsos com Amplitudes iguais a: <table border="1" style="margin: auto;"> <tr> <td>0,3 (P) + 0,8 (I)</td><td>0,6 (P) + 0,1 (I)</td> </tr> <tr> <td>0,3 (P) + 0,4 (I)</td><td>0,2 (P) + 0,6 (I)</td> </tr> <tr> <td>0,4 (P) + 0,9 (I)</td><td>0,5 (P) + 0,2 (I)</td> </tr> <tr> <td>0,7 (P) + 0,5 (I)</td><td>0,8 (P) + 0,3 (I)</td> </tr> </table>	0,3 (P) + 0,8 (I)	0,6 (P) + 0,1 (I)	0,3 (P) + 0,4 (I)	0,2 (P) + 0,6 (I)	0,4 (P) + 0,9 (I)	0,5 (P) + 0,2 (I)	0,7 (P) + 0,5 (I)	0,8 (P) + 0,3 (I)						
0,3 (P) + 0,8 (I)	0,6 (P) + 0,1 (I)													
0,3 (P) + 0,4 (I)	0,2 (P) + 0,6 (I)													
0,4 (P) + 0,9 (I)	0,5 (P) + 0,2 (I)													
0,7 (P) + 0,5 (I)	0,8 (P) + 0,3 (I)													

Resultados Esperados		Resultados Obtidos			
		BACKPROPAGATION		AG	
<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>
<i>Parábola</i>	<i>Impulso</i>	<i>Parábola</i>	<i>Impulso</i>	<i>Parábola</i>	<i>Impulso</i>
0,5	0,6	0,4629	0,6487	0,4687	0,6325
0,4	0,9	0,4000	0,9000	0,4099	0,6836
0,7	0,1	0,7025	0,1066	0,6704	0,1404
0,5	0,2	0,5000	0,2000	0,5350	0,3037

Tabela 5.7: Resultados da simulação para padrões como parábolas e impulsos

Nesta segunda simulação os resultados obtidos pelos dois métodos podem ser considerados como bons, porém a aproximação realizada pela RNA treinada pelo backpropagation foi melhor.

Vale lembrar que o treinamento pelo backpropagation depende muito da superfície de erro gerada ao longo do processo. O fato dos resultados produzidos pelo treinamento via AG não serem tão bons quanto os produzidos pelo Backpropagation se deve ao fato do AG ser um processo aleatório, o que não garante que sempre encontrará o melhor resultado.

Na próxima simulação são utilizadas parábolas e ondas triangulares, assim como composições das mesmas, para formar o conjunto de treinamento.

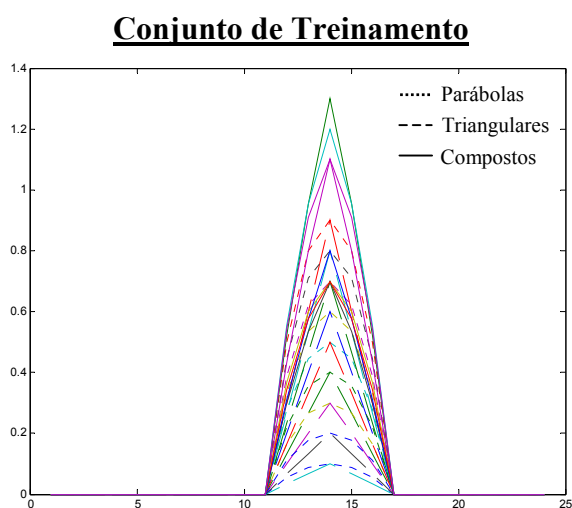


Figura 5.16: Conjunto de treinamento formado por parábolas e ondas triangulares, assim como composições destes

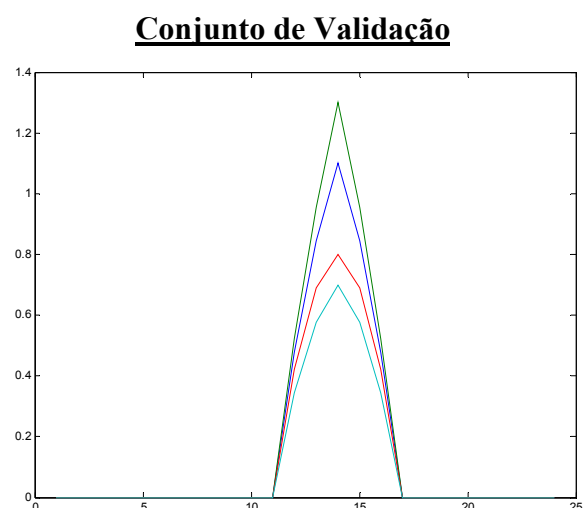


Figura 5.17: Conjunto de validação formado por composições de parábolas e ondas triangulares

Padrões Utilizados:																															
<u>Treinamento</u>	<u>Validação</u>																														
Parábolas com Amplitudes iguais a: <table border="1"> <tr> <td>0,1</td><td>0,2</td><td>0,3</td><td>0,4</td><td>0,5</td><td>0,6</td><td>0,7</td><td>0,8</td><td>0,9</td> </tr> </table> Onda Triangulares com Amplitudes iguais a: <table border="1"> <tr> <td>0,1</td><td>0,2</td><td>0,3</td><td>0,4</td><td>0,5</td><td>0,6</td><td>0,7</td><td>0,8</td><td>0,9</td> </tr> </table> Somas de Parábolas e Ondas Triangulares com Amplitudes iguais a: <table border="1"> <tr> <td>0,3 (P) + 0,8 (T)</td><td>0,6 (P) + 0,1 (T)</td> </tr> <tr> <td>0,3 (P) + 0,4 (T)</td><td>0,2 (P) + 0,6 (T)</td> </tr> <tr> <td>0,4 (P) + 0,9 (T)</td><td>0,5 (P) + 0,2 (T)</td> </tr> <tr> <td>0,7 (P) + 0,5 (T)</td><td>0,8 (P) + 0,3 (T)</td> </tr> </table>	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,3 (P) + 0,8 (T)	0,6 (P) + 0,1 (T)	0,3 (P) + 0,4 (T)	0,2 (P) + 0,6 (T)	0,4 (P) + 0,9 (T)	0,5 (P) + 0,2 (T)	0,7 (P) + 0,5 (T)	0,8 (P) + 0,3 (T)	<table border="1"> <tr> <td>0,5 (P) + 0,6 (T)</td> </tr> <tr> <td>0,4 (P) + 0,9 (T)</td> </tr> <tr> <td>0,7 (P) + 0,1 (T)</td> </tr> <tr> <td>0,5 (P) + 0,2 (T)</td> </tr> </table>	0,5 (P) + 0,6 (T)	0,4 (P) + 0,9 (T)	0,7 (P) + 0,1 (T)	0,5 (P) + 0,2 (T)
0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9																							
0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9																							
0,3 (P) + 0,8 (T)	0,6 (P) + 0,1 (T)																														
0,3 (P) + 0,4 (T)	0,2 (P) + 0,6 (T)																														
0,4 (P) + 0,9 (T)	0,5 (P) + 0,2 (T)																														
0,7 (P) + 0,5 (T)	0,8 (P) + 0,3 (T)																														
0,5 (P) + 0,6 (T)																															
0,4 (P) + 0,9 (T)																															
0,7 (P) + 0,1 (T)																															
0,5 (P) + 0,2 (T)																															

Resultados Esperados		Resultados Obtidos			
		BACKPROPAGATION		AG	
<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>
<i>Parábola</i>	<i>Onda Triangular</i>	<i>Parábola</i>	<i>Onda Triangular</i>	<i>Parábola</i>	<i>Onda Triangular</i>
0,5	0,6	0,4501	0,5291	0,5374	0,5016
0,4	0,9	0,4000	0,9000	0,4513	0,6894
0,7	0,1	0,6912	0,0947	0,6975	0,1151
0,5	0,2	0,5002	0,2001	0,5327	0,2335

Tabela 5.8: Resultados da simulação para padrões como parábolas e ondas triangulares

Assim como tem ocorrido até agora os resultados obtidos após a inclusão de padrões compostos no conjunto de treinamento tem sido muito próximos dos desejados, gerando assim valores de erro muito pequenos. Mais uma vez os melhores resultados são aqueles gerados pelo Backpropagation, porém os gerados pelo AG não são considerados desprezíveis.

A próxima simulação utiliza pulsos retangulares e ondas triangulares, assim composição destas para formar o conjunto de treinamento. Vários treinamentos foram realizados e os melhores resultados obtidos estão expostos a seguir.

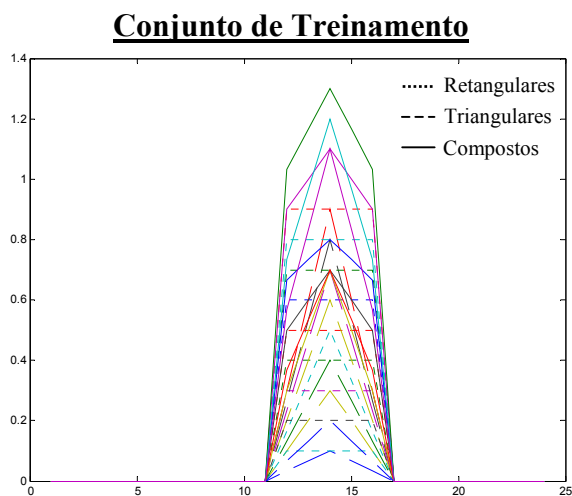


Figura 5.18: Conjunto de treinamento formado por ondas triangulares e pulsos retangulares, assim como composições destes

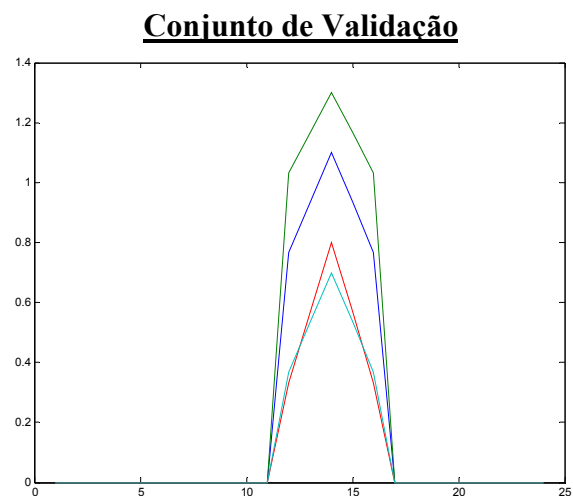


Figura 5.19: Conjunto de validação formado por composições de ondas triangulares e pulsos retangulares

Padrões Utilizados:																															
<u>Treinamento</u>	<u>Validação</u>																														
<p>Ondas Triangulares com Amplitudes iguais a:</p> <table border="1" style="width: 100%; text-align: center;"> <tr> <td>0,1</td><td>0,2</td><td>0,3</td><td>0,4</td><td>0,5</td><td>0,6</td><td>0,7</td><td>0,8</td><td>0,9</td> </tr> </table> <p>Pulsos Retangulares com Amplitudes iguais a:</p> <table border="1" style="width: 100%; text-align: center;"> <tr> <td>0,1</td><td>0,2</td><td>0,3</td><td>0,4</td><td>0,5</td><td>0,6</td><td>0,7</td><td>0,8</td><td>0,9</td> </tr> </table> <p>Somas Ondas Triangulares e Pulsos Retangulares com Amplitudes iguais a:</p> <table border="1" style="width: 100%; text-align: center;"> <tr> <td>0,3 (T) + 0,8 (R)</td><td>0,6 (T) + 0,1 (R)</td> </tr> <tr> <td>0,3 (T) + 0,4 (R)</td><td>0,2 (T) + 0,6 (R)</td> </tr> <tr> <td>0,4 (T) + 0,9 (R)</td><td>0,5 (T) + 0,2 (R)</td> </tr> <tr> <td>0,7 (T) + 0,5 (R)</td><td>0,8 (T) + 0,3 (R)</td> </tr> </table>	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,3 (T) + 0,8 (R)	0,6 (T) + 0,1 (R)	0,3 (T) + 0,4 (R)	0,2 (T) + 0,6 (R)	0,4 (T) + 0,9 (R)	0,5 (T) + 0,2 (R)	0,7 (T) + 0,5 (R)	0,8 (T) + 0,3 (R)	<table border="1" style="width: 100%; text-align: center;"> <tr> <td>0,5 (T) + 0,6 (R)</td> </tr> <tr> <td>0,4 (T) + 0,9 (R)</td> </tr> <tr> <td>0,7 (T) + 0,1 (R)</td> </tr> <tr> <td>0,5 (T) + 0,2 (R)</td> </tr> </table>	0,5 (T) + 0,6 (R)	0,4 (T) + 0,9 (R)	0,7 (T) + 0,1 (R)	0,5 (T) + 0,2 (R)
0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9																							
0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9																							
0,3 (T) + 0,8 (R)	0,6 (T) + 0,1 (R)																														
0,3 (T) + 0,4 (R)	0,2 (T) + 0,6 (R)																														
0,4 (T) + 0,9 (R)	0,5 (T) + 0,2 (R)																														
0,7 (T) + 0,5 (R)	0,8 (T) + 0,3 (R)																														
0,5 (T) + 0,6 (R)																															
0,4 (T) + 0,9 (R)																															
0,7 (T) + 0,1 (R)																															
0,5 (T) + 0,2 (R)																															

Resultados Esperados		Resultados Obtidos			
		BACKPROPAGATION		AG	
<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>
<i>Onda Triangular</i>	<i>Pulso Retangular</i>	<i>Onda Triangular</i>	<i>Pulso Retangular</i>	<i>Onda Triangular</i>	<i>Pulso Retangular</i>
0,5	0,6	0,4767	0,5942	0,6211	0,5903
0,4	0,9	0,4000	0,8999	0,4953	0,7444
0,7	0,1	0,7021	0,1024	0,5627	0,1735
0,5	0,2	0,5001	0,1988	0,5137	0,2268

Tabela 5.9: Resultados da simulação para padrões como ondas triangulares e pulsos retangulares

Pelos resultados apresentados acima, é possível perceber que neste último caso àqueles referentes ao treinamento realizado com AG não são tão bons quanto os anteriores e que os obtidos pelo backpropagation estão muito próximos dos desejados.

Isto nos permite a primeira conclusão: nos casos onde há uma relativa simplicidade dos padrões, a superfície de erro gerada tende a ser simples e unimodal. Sendo assim, não há necessidade de se utilizar AG, pois custo computacional é elevado e não há risco do ocorrer convergência prematura devido ao estacionamento do vetor gradiente em um ponto de mínimo local.

Após a realização das simulações envolvendo padrões naturais simples e compostos, passaremos a considerar como padrões de treinamento curvas de carga referentes ao consumo de energia elétrica de três diferentes tipos de consumidores: residencial, industrial e comercial. No próximo tópico serão dados detalhes de todas as simulações realizadas com estes novos padrões assim como os resultados obtidos.

5.5 DEFINIÇÃO DOS NOVOS PADRÕES DE TREINAMENTO

O primeiro passo neste capítulo é a definição das curvas de carga típicas que ilustram o comportamento dos diversos tipos de setores consumidores de energia elétrica. A aplicação prática proposta sugere que, a partir de uma curva de carga referente a um ponto qualquer do sistema de distribuição de energia elétrica, possa se determinar a porção relativa a cada um destes setores de acordo com o comportamento típico de cada um. A cada uma destas curvas típicas associamos um comportamento tido como padrão para aquele setor consumidor, desta forma estamos determinando curvas que serão padrão para os setores residencial, comercial e industrial. Os gráficos abaixo apresentam estes novos padrões que serão usados para as simulações seguintes.

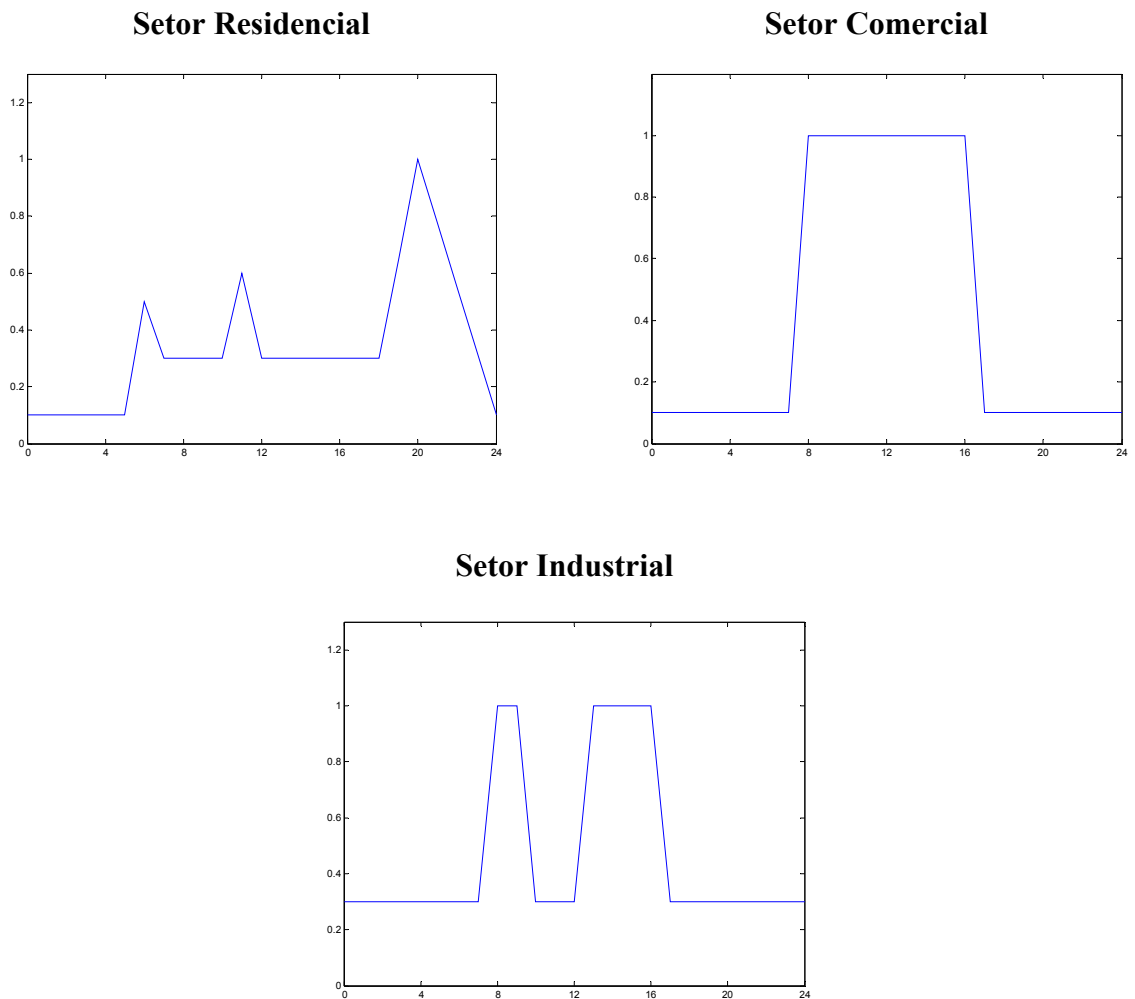


Figura 5.20: Curvas adotadas como referências aos Setores Residencial, Comercial e Industrial

A idéia de reconhecer padrões contidos em curvas é uma aplicação característica de RNA, cujo ponto forte é o reconhecimento de padrões. Sendo assim, sugere-se a implementação de um sistema capaz de realizar tal operação. Este sistema tanto pode ser uma RNA simples como uma hibridização entre RNA e AG. Os resultados apresentados a seguir são produzidos por estes dois tipos. Os conjuntos utilizados para o treinamento das RNA são os mesmos para os dois casos e a configuração dos parâmetros é apresentada ao final das simulações.

Definidos os padrões, o passo seguinte é a criação dos conjuntos de treinamento, a definição dos parâmetros da RNA, do AG e a realização dos testes de avaliação da performance. Para isto é necessário também um conjunto de validação, apresentado após os dados referentes ao conjunto de treinamento.

O primeiro padrão a ser verificado é o residencial e é feita apenas confirmação da RNA quanto à capacidade de identificação do padrão.

Para o treinamento são utilizadas seis amostras da curva típica do setor residencial com diferentes amplitudes, e quatro novas amostras com amplitudes diferentes daquelas usadas anteriormente são usadas para a validação.

Os detalhes desta primeira simulação são apresentados a seguir.

Conjunto de Treinamento

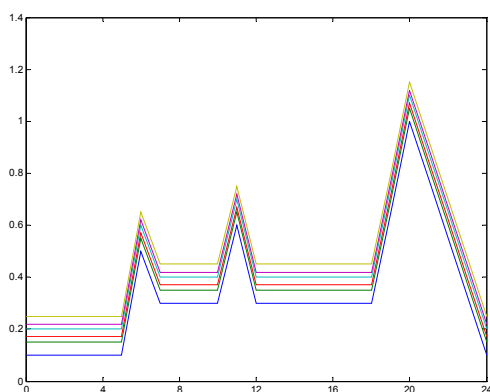


Figura 5.21: Conjunto de treinamento formado por curvas representativas do setor residencial, com várias amplitudes

Conjunto de Validação

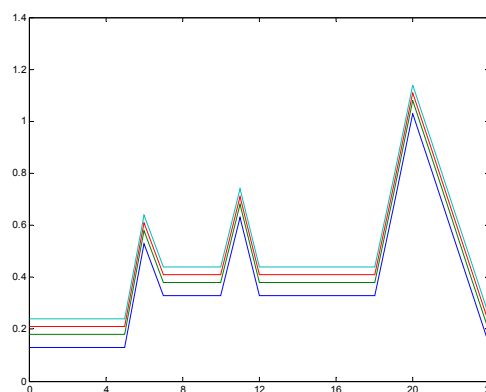


Figura 5.22: Conjunto de validação formado por curvas do setor residencial, com amplitudes diferentes das utilizadas para o treinamento

<i>Padrões Utilizados:</i>											
<i>Treinamento</i>	<i>Validação</i>										
Curvas de Carga do Setor Residencial com valores iniciais de Amplitude iguais a:	Curvas de Carga do Setor Residencial com valores iniciais de Amplitude iguais a:										
<table border="1"> <tr> <td>0,10</td> <td>0,15</td> <td>0,17</td> <td>0,20</td> <td>0,22</td> <td>0,25</td> </tr> </table>	0,10	0,15	0,17	0,20	0,22	0,25	<table border="1"> <tr> <td>0,13</td> <td>0,18</td> <td>0,21</td> <td>0,24</td> </tr> </table>	0,13	0,18	0,21	0,24
0,10	0,15	0,17	0,20	0,22	0,25						
0,13	0,18	0,21	0,24								

Resultados Esperados	Resultados Obtidos	
	<i>Backpropagation</i>	<i>AG</i>
Saída	Saída	Saída
0,13	0,1307	0,1274
0,18	0,1801	0,1761
0,21	0,2010	0,2107
0,24	0,2402	0,2478
SSE:	8.15e-005	8.33e-005

Tabela 5.10: Resultados da simulação realizada com o padrão residencial

O erro calculado acima corresponde ao somatório dos erros quadráticos que é utilizado para avaliar a performance das RNA.

Neste primeiro treinamento pode-se notar, que apesar de sua simplicidade, os dois métodos apresentaram bons resultados, havendo uma aproximação melhor através da RNA treinada pelo backpropagation.

A próxima simulação é realizada da mesma forma, porém com a curva referente ao setor industrial. Neste caso os valores iniciais de amplitude são diferentes dos anteriores.

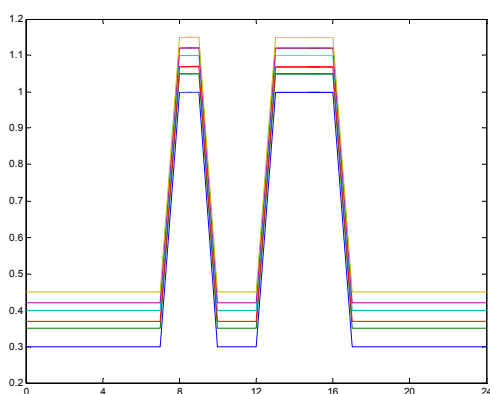


Figura 5.23: Conjunto de treinamento formado por curvas representativas do setor industrial, com várias amplitudes

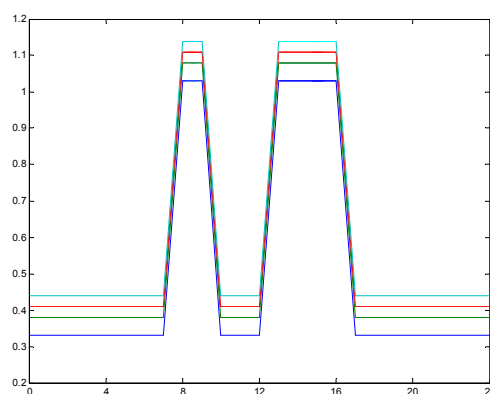


Figura 5.24: Conjunto de validação formado por curvas do setor industrial, com amplitudes diferentes das utilizadas para o treinamento

Padrões Utilizados:											
<u>Treinamento</u>	<u>Validação</u>										
Curvas de Carga do Setor Industrial com valores iniciais de Amplitude iguais a:	Curvas de Carga do Setor Industrial com valores iniciais de Amplitude iguais a:										
<table border="1"> <tr> <td>0,30</td> <td>0,35</td> <td>0,37</td> <td>0,40</td> <td>0,42</td> <td>0,45</td> </tr> </table>	0,30	0,35	0,37	0,40	0,42	0,45	<table border="1"> <tr> <td>0,33</td> <td>0,38</td> <td>0,41</td> <td>0,44</td> </tr> </table>	0,33	0,38	0,41	0,44
0,30	0,35	0,37	0,40	0,42	0,45						
0,33	0,38	0,41	0,44								

Resultados Esperados	Resultados Obtidos	
	<i>Backpropagation</i>	<i>AG</i>
Saída	Saída	Saída
0,33	0,3317	0,3271
0,38	0,3807	0,3754
0,41	0,4093	0,4097
0,44	0,4417	0,4448
SSE:	6.76e-006	5.27e-005

Tabela 5.11: Resultados da simulação realizada com o padrão industrial

Neste último caso as pode-se considerar os resultados muito próximos aos esperados, em ambas as RNA. Para verificação final da condição de reconhecimento de padrões será simulado o caso onde o padrão de treinamento passa a ser a curva do setor comercial, conforme mostrado a seguir.

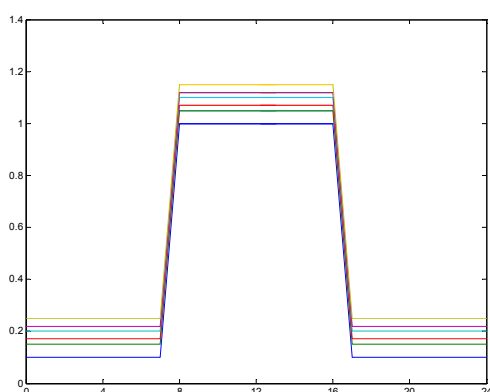


Figura 5.25: Conjunto de treinamento formado por curvas representativas do setor comercial, com várias amplitudes

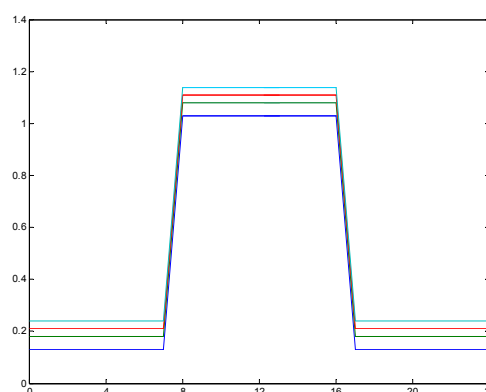


Figura 5.26: Conjunto de treinamento formado por curvas do setor comercial com amplitudes diferentes das utilizadas para o treinamento

<i>Padrões Utilizados:</i>	
<u><i>Treinamento</i></u>	<u><i>Validação</i></u>
Curvas de Carga do Setor Comercial com valores iniciais de Amplitude iguais a:	Curvas de Carga do Setor Comercial com valores iniciais de Amplitude iguais a:
0,10 0,15 0,17 0,20 0,22 0,25	0,13 0,18 0,21 0,24

Resultados Esperados	Resultados Obtidos	
	<i>Backpropagation</i>	<i>AG</i>
Saída	Saída	Saída
0,13	0,1302	0,1322
0,18	0,1799	0,1841
0,21	0,2100	0,2091
0,24	0,2399	0,2302
<i>SSE:</i>	<i>6.00e-008</i>	<i>1.185e-004</i>

Tabela 5.12: Resultados da simulação realizada com o padrão comercial

Neste último caso o treinamento pode ser considerado um tanto quanto simples e conforme esperado as RNA responderam de acordo com o desejado.

De posse dos resultados obtidos e apresentados acima, é importante concluir e reafirmar que não houve diferença expressiva entre os resultados gerados pelo sistema híbrido proposto e a RNA treinada da forma convencional.

Com relação ao custo computacional pode-se dizer que o treinamento realizado pelo backpropagation supera com grande vantagem o AG.

Para se ter idéia, nos treinamentos realizados neste trabalho, o tempo gasto pelo AG ao treinar uma RNA foi cerca de 30 a 50 vezes maior que o gasto pelo backpropagation.

Conforme dito, foram utilizadas toolboxes do MATLAB para a execução das simulações, porém o software as oferece separadamente, não permitindo, na toolbox de RNA, o treinamento via AG.

Sendo assim foi necessária uma adaptação, ou seja, para o sistema híbrido utilizou-se a toolbox de AG, e o erro gerado pela RNA como função objetivo.

Todo o código referente à RNA teve que ser escrito em linguagem proprietária do MATLAB e o mesmo está disponibilizado em Anexo.

A partir deste ponto são realizados treinamentos mais complexos, onde a verificação das RNA é feita através da composição dos padrões de entrada, assim como foi feito anteriormente com os padrões naturais.

5.5.1 Simulações envolvendo composições entre os novos padrões

Neste tópico serão apresentadas as simulações envolvendo composições dos padrões apresentados no item anterior. A idéia de compor padrões se justifica, pois a intenção é avaliar uma curva de carga adquirida em um ponto qualquer da rede elétrica de distribuição e esta nada mais será do que a composição de curvas de cargas dos setores consumidores presentes naquele ponto do sistema. Sendo assim passaremos a avaliar o desempenho das RNA para as situações onde há a composição de curvas.

Baseado nas experiências adquiridas anteriormente nos treinamentos com padrões simples e testes com padrões compostos, os novos conjuntos de treinamento terão a inclusão de alguns padrões compostos, na intenção de “ensinar” a RNA lidar com estas composições.

A primeira simulação envolve curvas referentes ao setor residencial e comercial, assim como composições das mesmas. Os dados referentes a esta simulação estão apresentados a seguir:

Conjunto de Treinamento

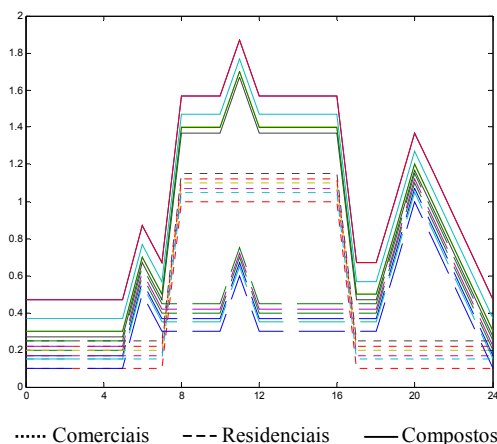


Figura 5.27: Conjunto de treinamento formado por curvas representativas dos setores residencial e comercial, assim como composições das mesmas

Conjunto de Validação

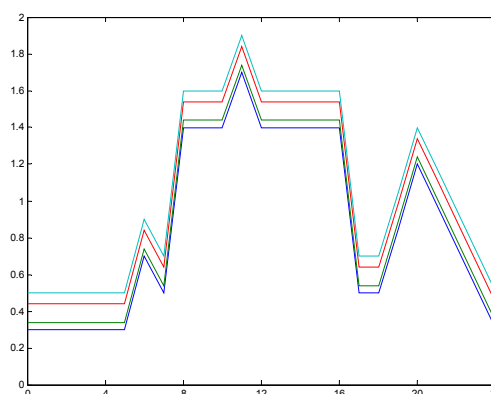


Figura 5.28: Conjunto de validação formado por composições de curvas representativas dos setores residencial e comercial

Padrões Utilizados:											
<u>Treinamento</u>	<u>Validação</u>										
Amplitudes iniciais do Setor Residencial:	Composições de curvas referentes ao Setor Residencial e Comercial com os seguintes valores iniciais de Amplitude:										
<table border="1" style="margin: auto;"> <tr> <td>0,17</td><td>0,20</td><td>0,15</td><td>0,1</td><td>0,25</td><td>0,22</td> </tr> </table>		0,17	0,20	0,15	0,1	0,25	0,22				
0,17		0,20	0,15	0,1	0,25	0,22					
Amplitudes iniciais do Setor Comercial:											
<table border="1" style="margin: auto;"> <tr> <td>0,10</td><td>0,17</td><td>0,20</td><td>0,25</td><td>0,22</td><td>0,15</td> </tr> </table>	0,10	0,17	0,20	0,25	0,22	0,15					
0,10	0,17	0,20	0,25	0,22	0,15						
Composição das anteriores com as seguintes amplitudes iniciais:											
<table border="1" style="margin: auto;"> <tr> <td>0,10 (R) + 0,20 (C)</td><td>0,15 (R) + 0,10 (C)</td> </tr> <tr> <td>0,17 (R) + 0,10 (C)</td><td>0,22 (R) + 0,25 (C)</td> </tr> <tr> <td>0,25 (R) + 0,22 (C)</td><td>0,20 (R) + 0,17 (C)</td> </tr> </table>	0,10 (R) + 0,20 (C)	0,15 (R) + 0,10 (C)	0,17 (R) + 0,10 (C)	0,22 (R) + 0,25 (C)	0,25 (R) + 0,22 (C)	0,20 (R) + 0,17 (C)	<table border="1" style="margin: auto;"> <tr> <td>0,15 (R) + 0,15 (C)</td> </tr> <tr> <td>0,17 (R) + 0,17 (C)</td> </tr> <tr> <td>0,22 (R) + 0,22 (C)</td> </tr> <tr> <td>0,25 (R) + 0,25 (C)</td> </tr> </table>	0,15 (R) + 0,15 (C)	0,17 (R) + 0,17 (C)	0,22 (R) + 0,22 (C)	0,25 (R) + 0,25 (C)
0,10 (R) + 0,20 (C)	0,15 (R) + 0,10 (C)										
0,17 (R) + 0,10 (C)	0,22 (R) + 0,25 (C)										
0,25 (R) + 0,22 (C)	0,20 (R) + 0,17 (C)										
0,15 (R) + 0,15 (C)											
0,17 (R) + 0,17 (C)											
0,22 (R) + 0,22 (C)											
0,25 (R) + 0,25 (C)											
<p><i>Onde:</i> 0.10 (R) + 0.20 (C) representa a composição formada pela onda referente ao setor residencial com amplitude inicial igual a 0.10 somada a onda referente ao setor comercial com amplitude inicial igual a 0.20</p>											

É importante ressaltar que no conjunto de treinamento são utilizadas doze amostras de padrões simples e seis composições destes, o que é suficiente para o aprendizado da RNA

Resultados Esperados		Resultados Obtidos			
		BACKPROPAGATION		AG	
<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>
<i>Residencial</i>	<i>Comercial</i>	<i>Residencial</i>	<i>Comercial</i>	<i>Residencial</i>	<i>Comercial</i>
0,15	0,15	0,1502	0,1501	0,1514	0,1503
0,17	0,17	0,1709	0,1708	0,1629	0,1583
0,22	0,22	0,2204	0,2205	0,2092	0,1908
0,25	0,25	0,2484	0,2488	0,2590	0,2251
SSE:		6.00e-006		0.0019	

Tabela 5.13: Resultados da simulação realizada com os padrões residencial e comercial

Novamente, para avaliação da performance das RNA utilizamos a soma dos erros quadráticos produzidos pela diferença entre os valores desejados na saída da RNA e os obtidos nas simulações.

De acordo com os resultados apresentados acima, a primeira simulação foi bem sucedida nos dois casos. A diferença entre os erros comprova a afirmação anterior.

Embora o treinamento realizado via AG consuma um tempo razoavelmente maior para encontrar o conjunto de pesos ideal, sua utilização pode ser uma boa alternativa nos casos onde o treinamento pelo método convencional não gerar resultados satisfatórios.

Na próxima simulação foram utilizados como padrões de treinamento amostras das curvas referentes aos setores residencial e industrial, assim como composições destas.

Os dados referentes a esta simulação estão apresentados a seguir.

Conjunto de Treinamento

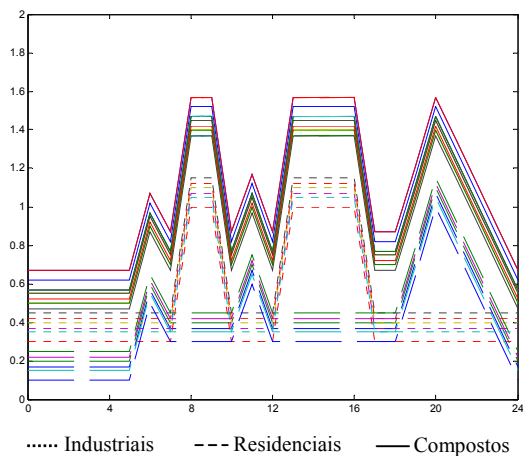


Figura 5.29: Conjunto de treinamento formado por curvas representativas dos setores residencial e industrial, assim como composições das mesmas

Conjunto de Validação

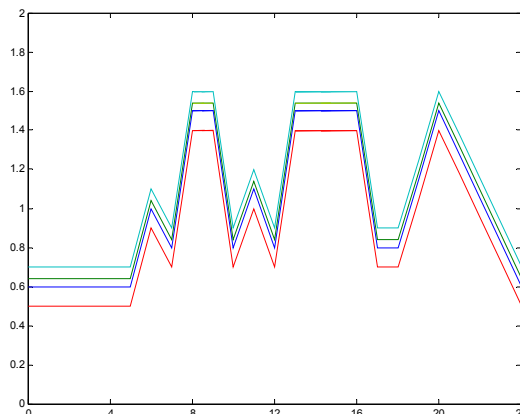


Figura 5.30: Conjunto de validação formado por composições de curvas representativas dos setores residencial e industrial

Padrões Utilizados:

Treinamento

Amplitudes iniciais do Setor Residencial:

0,17	0,20	0,15	0,1	0,25	0,22
------	------	------	-----	------	------

Amplitudes iniciais do Setor Industrial:

0,30	0,37	0,40	0,45	0,42	0,35
------	------	------	------	------	------

Composição das anteriores com as seguintes amplitudes iniciais:

0,10 (R) + 0,40 (I)	0,15 (R) + 0,40 (I)
0,17 (R) + 0,30 (I)	0,22 (R) + 0,45 (I)
0,25 (R) + 0,42 (I)	0,20 (R) + 0,37 (I)

Onde: 0,10 (R) + 0,20 (I) representa a composição formada pela onda referente ao setor residencial com amplitude inicial igual a 0,10 somada a onda referente ao setor industrial com amplitude inicial igual a 0,20

Validação

Composições de curvas referentes ao Setor Residencial e Industrial com os seguintes valores iniciais de Amplitude:

0,20 (R) + 0,40 (I)
0,22 (R) + 0,42 (I)
0,15 (R) + 0,35 (I)
0,25 (R) + 0,45 (I)

Resultados Esperados		Resultados Obtidos			
		BACKPROPAGATION		AG	
<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>
<i>Residencial</i>	<i>Industrial</i>	<i>Residencial</i>	<i>Industrial</i>	<i>Residencial</i>	<i>Industrial</i>
0,20	0,40	0,2020	0,4084	0,2066	0,3973
0,22	0,42	0,2209	0,4237	0,2145	0,4182
0,15	0,35	0,1496	0,3524	0,1854	0,3413
0,25	0,45	0,2460	0,4410	0,2258	0,4476
SSE:		0.00019038		0.0020	

Tabela 5.14: Resultados da simulação realizada com os padrões residencial e industrial

Os resultados acima novamente revelam a viabilidade de se utilizar AG como alternativa ao backpropagation. A diferença entre os valores de erro é de 0.0018, o que significa que em termos gerais os resultados, em ambos os casos, estão muito próximos dos desejados. Para a próxima simulação foram escolhidos os setores industrial e comercial.

O conjunto de treinamento é formado de acordo com os critérios adotados nos casos anteriores. O conjunto de validação é formado por amostras não utilizadas para o treinamento, o que exprime a capacidade da RNA em aprender.

Conjunto de Treinamento

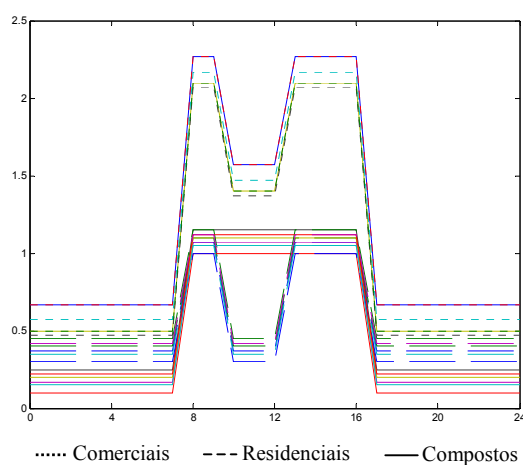


Figura 5.31: Conjunto de treinamento formado por curvas representativas dos setores industrial e comercial, assim como composições das mesmas

Conjunto de Validação

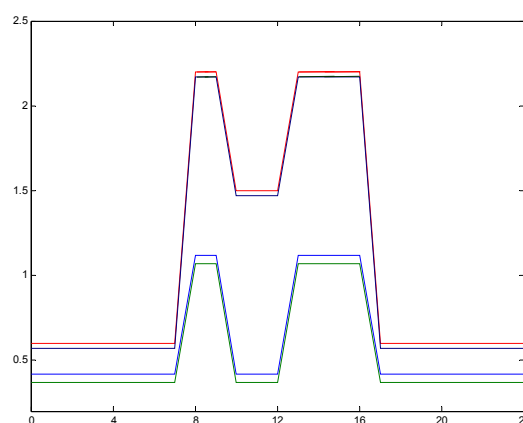


Figura 5.32: Conjunto de validação formado por composições de curvas representativas dos setores industrial e comercial

Padrões Utilizados:							
<u>Treinamento</u>	<u>Validação</u>						
Amplitudes iniciais do Setor Industrial:	Composições de curvas referentes ao Setor Industrial e Comercial com os seguintes valores iniciais de Amplitude:						
<table border="1"><tr><td>0,30</td><td>0,37</td><td>0,40</td><td>0,45</td><td>0,42</td><td>0,35</td></tr></table>		0,30	0,37	0,40	0,45	0,42	0,35
0,30		0,37	0,40	0,45	0,42	0,35	
Amplitudes iniciais do Setor Comercial:							
<table border="1"><tr><td>0,17</td><td>0,20</td><td>0,15</td><td>0,1</td><td>0,25</td><td>0,22</td></tr></table>	0,17	0,20	0,15	0,1	0,25	0,22	
0,17	0,20	0,15	0,1	0,25	0,22		
Composição das anteriores com as seguintes amplitudes iniciais:	<table border="1"><tr><td>0,42 (I)</td></tr><tr><td>0,37 (I)</td></tr><tr><td>0,40 (I) + 0,20 (C)</td></tr><tr><td>0,40 (I) + 0,17 (C)</td></tr></table>	0,42 (I)	0,37 (I)	0,40 (I) + 0,20 (C)	0,40 (I) + 0,17 (C)		
0,42 (I)							
0,37 (I)							
0,40 (I) + 0,20 (C)							
0,40 (I) + 0,17 (C)							
<table border="1"><tr><td>0,30 (I) + 0,20 (C)</td><td>0,35 (I) + 0,15 (C)</td></tr><tr><td>0,37 (I) + 0,10 (C)</td><td>0,42 (I) + 0,25 (C)</td></tr><tr><td>0,45 (I) + 0,22 (C)</td><td>0,45 (I) + 0,17 (C)</td></tr></table>	0,30 (I) + 0,20 (C)	0,35 (I) + 0,15 (C)	0,37 (I) + 0,10 (C)	0,42 (I) + 0,25 (C)	0,45 (I) + 0,22 (C)	0,45 (I) + 0,17 (C)	
0,30 (I) + 0,20 (C)	0,35 (I) + 0,15 (C)						
0,37 (I) + 0,10 (C)	0,42 (I) + 0,25 (C)						
0,45 (I) + 0,22 (C)	0,45 (I) + 0,17 (C)						

Resultados Esperados		Resultados Obtidos			
		BACKPROPAGATION		AG	
<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>
<i>Industrial</i>	<i>Comercial</i>	<i>Industrial</i>	<i>Comercial</i>	<i>Industrial</i>	<i>Comercial</i>
0,42	0	0,4204	0,0002	0,3784	0,0396
0,37	0	0,3701	0,0002	0,3650	0,0274
0,40	0,20	0,4015	0,2011	0,4025	0,1963
0,40	0,17	0,3866	0,1860	0,3915	0,1896
SSE:		0.00043915		0.0045	

Tabela 5.15: Resultados da simulação realizada com os padrões industrial e comercial

Esta última simulação apresentou, como as anteriores, bons resultados nos dois casos. Apesar da diferença entre os erros ter sido pouco maior que nos casos anteriores (0.0041), pode-se considerar como bom o desempenho do AG no reconhecimento de padrões conjugados.

Terminada esta segunda etapa de simulações tem-se as seguintes conclusões:

- O sucesso no treinamento de uma RNA é fortemente dependente da consistência do conjunto de treinamento.

- A topologias utilizadas foram as mesmas em todos os casos, ou seja, a performance dos algoritmos utilizados para o treinamento não foi afetada por este parâmetro.
- O AG necessita de um número relativamente maior de gerações para a convergência.
- O sucesso no treinamento via AG depende muito do tamanho do espaço de busca. Em casos onde o intervalo que delimita o espaço de busca foi menor a convergência tendeu a ser mais rápida.
- O número de indivíduos no AG afeta fortemente seu desempenho. Nos casos onde houve um número muito grande ou muito pequeno de indivíduos na população o AG não encontrou resultados satisfatórios. Populações muito numerosas podem fazer com que o algoritmo não consiga encontrar o indivíduo ideal e populações pouco numerosas podem fazer com que o melhor não esteja presente.

Terminadas as conclusões o trabalho segue com a adição de uma nova característica aos padrões até então utilizados: a possibilidade de oscilações em seus valores de amplitude. Este será o foco do próximo item, a consideração nas oscilações decorrentes de variações no consumo.

5.6 CONSIDERAÇÕES SOBRE OSCILAÇÕES

O consumo de energia elétrica é algo extremamente variável com o tempo, e depende de uma enorme quantidade de fatores. Num dado momento o simples desligamento ou religação do maquinário de uma grande empresa resultará em mudanças na curva de carga da estação que fornece energia àquele trecho da rede de distribuição. Desta forma devem ser consideradas oscilações e o comportamento das RNA será avaliado nestas situações.

Um ponto importante com relação às oscilações no consumo é quanto à RNA, que pode assumir o comportamento de um filtro e simplesmente ignorar estas variações, tratando-as como um ruído.

As seguintes simulações utilizam-se o mesmo critério adotado para o treinamento das RNA anteriores, ou seja, se desejamos que a rede identifique oscilações nos padrões, devemos incorporar ao conjunto de treinamento algum tipo de informação referente. Para isso usamos amostras das curvas típicas dos setores considerados com algumas variações aleatórias de amplitude. Os conjuntos de validação também passarão a conter padrões com oscilações variando de 1 a 10% sobre os valores das amplitudes das curvas.

Seguindo a seqüência adotada anteriormente, iniciaremos esta fase de simulações com curvas típicas dos setores residencial e comercial.

Conjunto de Treinamento

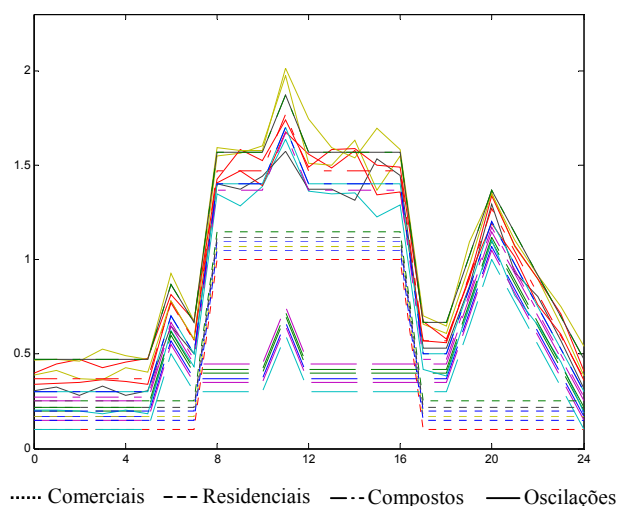


Figura 5.33: Conjunto de treinamento formado por curvas representativas dos setores residencial e comercial, assim como composições das mesmas submetidas a oscilações

Padrões Utilizados:

Amplitudes iniciais do Setor Residencial:

0,17	0,20	0,15	0,1	0,25	0,22
------	------	------	-----	------	------

Amplitudes iniciais do Setor Comercial:

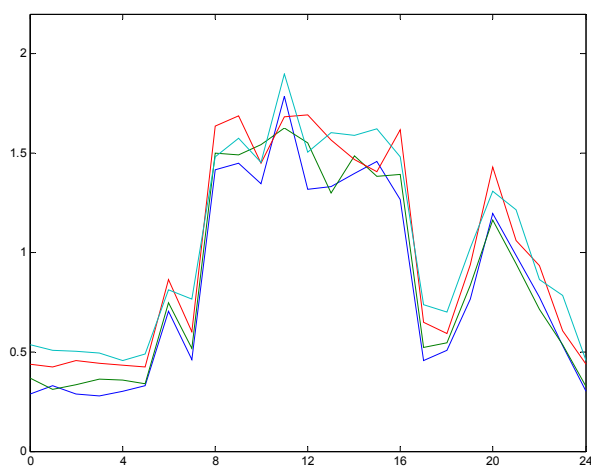
0,10	0,17	0,20	0,25	0,22	0,15
------	------	------	------	------	------

Composição das anteriores com as seguintes amplitudes iniciais:

0,10 (R) + 0,20 (C)	0,15 (R) + 0,10 (C)
0,17 (R) + 0,10 (C)	0,22 (R) + 0,25 (C)
0,25 (R) + 0,22 (C)	0,20 (R) + 0,17 (C)

* Sobre estes valores de amplitude considerou-se uma variação aleatória de 0 a 10%.

Conjunto de Validação



Padrões Utilizados:

Composições de curvas referentes ao Setor Residencial e Comercial com os seguintes valores iniciais de Amplitude:

$0,15 \pm 10\%$ (R) + $0,15 \pm 10\%$ (C)
$0,17 \pm 10\%$ (R) + $0,17 \pm 10\%$ (C)
$0,22 \pm 10\%$ (R) + $0,22 \pm 10\%$ (C)
$0,25 \pm 10\%$ (R) + $0,25 \pm 10\%$ (C)

Figura 5.34: Conjunto de validação formado por composições de curvas representativas dos setores residencial e comercial submetidas a oscilações

É importante notar que simples variações da ordem de 10% sobre os valores de amplitude provocam uma certa descaracterização nas curvas.

Primeiramente foram realizadas validações das RNA com oscilações da ordem de 10%, depois foram verificados os comportamentos das mesmas redes treinadas com amostras variando até 10% em amplitude para padrões com variações máximas de 5%.

Conjunto de Validação com oscilações máximas de $\pm 10\%$ dos valores de amplitude.

Resultados Esperados		Resultados Obtidos			
		BACKPROPAGATION		AG	
<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>
<i>Residencial</i>	<i>Comercial</i>	<i>Residencial</i>	<i>Comercial</i>	<i>Residencial</i>	<i>Comercial</i>
$0,15 \pm 10\%$	$0,15 \pm 10\%$	0,1713	0,1709	0,1567	0,1628
$0,17 \pm 10\%$	$0,17 \pm 10\%$	0,1819	0,1856	0,1662	0,1711
$0,22 \pm 10\%$	$0,22 \pm 10\%$	0,2041	0,2052	0,1871	0,2349
$0,25 \pm 10\%$	$0,25 \pm 10\%$	0,2271	0,2278	0,1887	0,2644
SSE:		0.0028		0.0055	

Tabela 5.16: Resultados da simulação realizada com 10% de oscilações sobre as amplitudes dos padrões residencial e comercial

Conjunto de Validação com oscilações máximas de $\pm 5\%$ dos valores de amplitude.

Resultados Esperados		Resultados Obtidos			
		BACKPROPAGATION		AG	
<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>
<i>Residencial</i>	<i>Comercial</i>	<i>Residencial</i>	<i>Comercial</i>	<i>Residencial</i>	<i>Comercial</i>
0,15 \pm 5%	0,15 \pm 5%	0,1316	0,1347	0,1614	0,1401
0,17 \pm 5%	0,17 \pm 5%	0,1556	0,1588	0,1751	0,1866
0,22 \pm 5%	0,22 \pm 5%	0,2336	0,2333	0,1837	0,2053
0,25 \pm 5%	0,25 \pm 5%	0,2362	0,2356	0,1859	0,2545
SSE:		0.0017		0.0062	

Tabela 5.17: Resultados da simulação realizada com 5% de oscilações sobre as amplitudes dos padrões residencial e comercial

Neste primeiro caso os resultados apresentados pelas duas RNA foram satisfatórios. A primeira RNA treinada pelo backpropagation teve um desempenho pouco melhor, mas de qualquer forma o erro produzido pela segunda pode ser considerado relativamente baixo.

No caso seguinte serão avaliados os setores industrial e comercial. A forma de composição dos conjuntos de treinamento e validação será a mesma.

Conjunto de Treinamento

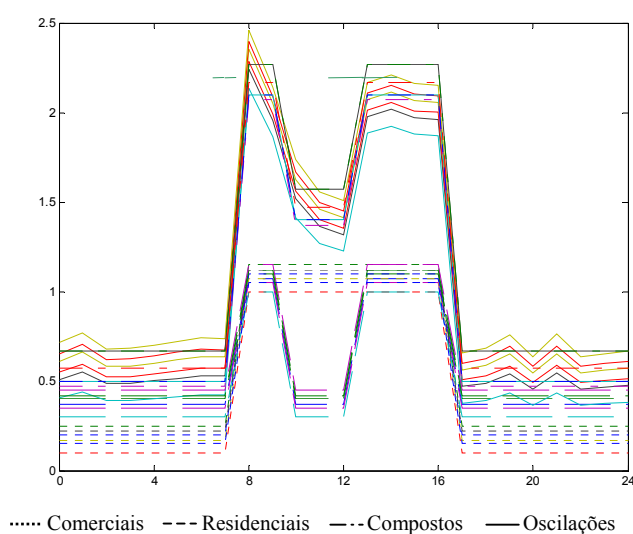


Figura 5.35: Conjunto de treinamento formado por curvas representativas dos setores industrial e comercial, assim como composições das mesmas submetidas a oscilações

Padrões Utilizados:

Amplitudes iniciais do Setor Industrial:

0,30	0,37	0,40	0,45	0,42	0,35
------	------	------	------	------	------

Amplitudes iniciais do Setor Comercial:

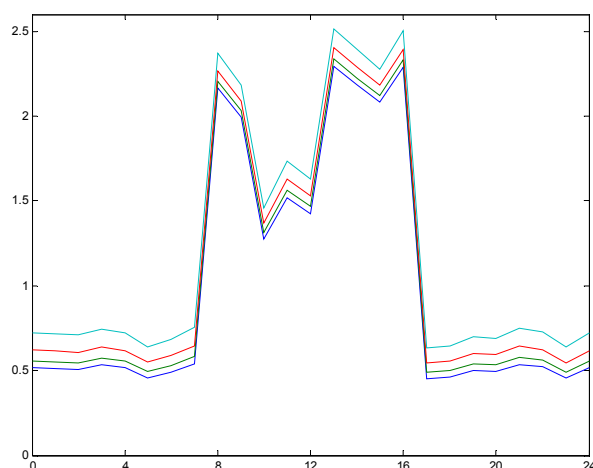
0,17	0,20	0,15	0,1	0,25	0,22
------	------	------	-----	------	------

Composição das anteriores com as seguintes amplitudes iniciais:

0,30 (I) + 0,20 (C)	0,35 (I) + 0,15 (C)
0,37 (I) + 0,10 (C)	0,42 (I) + 0,25 (C)
0,45 (I) + 0,22 (C)	0,45 (I) + 0,17 (C)

* Sobre estes valores de amplitude considerou-se uma variação aleatória de 0 a 10%.

Conjunto de Validação



Padrões Utilizados:

Composições de curvas referentes ao Setor Industrial e Comercial com os seguintes valores iniciais de Amplitude:

$0,30 \pm 10\%$ (I) + $0,20 \pm 10\%$ (C)
$0,37 \pm 10\%$ (I) + $0,17 \pm 10\%$ (C)
$0,40 \pm 10\%$ (I) + $0,20 \pm 10\%$ (C)
$0,40 \pm 10\%$ (I) + $0,14 \pm 10\%$ (C)

Figura 5.36: Conjunto de validação formado por composições de curvas representativas dos setores industrial e comercial submetidas a oscilações

Conjunto de Validação com oscilações máximas de $\pm 10\%$ dos valores de amplitude.

Resultados Esperados		Resultados Obtidos			
		BACKPROPAGATION		AG	
<u>Saída 1</u> <i>Industrial</i>	<u>Saída 2</u> <i>Comercial</i>	<u>Saída 1</u> <i>Industrial</i>	<u>Saída 2</u> <i>Comercial</i>	<u>Saída 1</u> <i>Industrial</i>	<u>Saída 2</u> <i>Comercial</i>
$0,30 \pm 10\%$	$0,20 \pm 10\%$	0,3443	0,2146	0,3278	0,1399
$0,37 \pm 10\%$	$0,17 \pm 10\%$	0,3640	0,2329	0,3614	0,1648
$0,40 \pm 10\%$	$0,20 \pm 10\%$	0,3953	0,2568	0,3932	0,1967
$0,40 \pm 10\%$	$0,14 \pm 10\%$	0,4554	0,2896	0,4158	0,2303
SSE:		0.0349		0.0129	

Tabela 5.18: Resultados da simulação realizada com 10% de oscilações sobre as amplitudes dos padrões industrial e comercial

Conjunto de Validação com oscilações máximas de $\pm 5\%$ dos valores de amplitude.

Resultados Esperados		Resultados Obtidos			
		BACKPROPAGATION		AG	
<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>	<u>Saída 1</u>	<u>Saída 2</u>
<i>Industrial</i>	<i>Comercial</i>	<i>Industrial</i>	<i>Comercial</i>	<i>Industrial</i>	<i>Comercial</i>
0,30 \pm 5%	0,20 \pm 5%	0,3566	0,1517	0,3228	0,2024
0,37 \pm 5%	0,17 \pm 5%	0,3749	0,1787	0,3539	0,2118
0,40 \pm 5%	0,20 \pm 5%	0,4050	0,2136	0,3844	0,2281
0,40 \pm 5%	0,14 \pm 5%	0,4655	0,2607	0,4081	0,2475
SSE:		0.0247		0.0152	

Tabela 5.19: Resultados da simulação realizada com 5% de oscilações sobre as amplitudes dos padrões industrial e comercial

De acordo com os resultados acima, podemos perceber a viabilidade do uso de AG em casos onde o treinamento convencional não encontra um conjunto de pesos ótimos para a RNA. Esta é uma das justificativas do uso de sistemas híbridos em determinados casos. Segundo [11] os “testes conduzidos com problemas simples de reconhecimento de padrões produziram resultados inconclusivos”. Isto mostra a especificidade de cada sistema, cabendo ao projetista a escolha do melhor algoritmo de treinamento.

A próxima simulação será realizada com curvas representativas dos setores residencial e industrial. Será também mais uma chance de comprovar a viabilidade do sistema híbrido proposto. A seguir são dados os detalhes desta simulação.

Conjunto de Treinamento

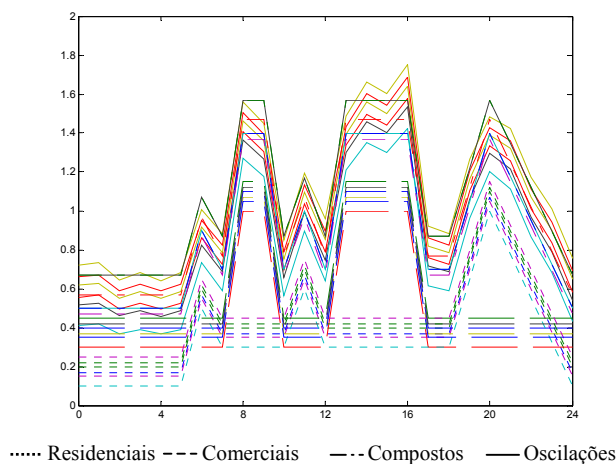


Figura 5.37: Conjunto de treinamento formado por curvas representativas dos setores residencial e industrial, assim como composições das mesmas submetidas a oscilações

Padrões Utilizados:

Amplitudes iniciais do Setor Residencial:

0,17	0,20	0,15	0,1	0,25	0,22
------	------	------	-----	------	------

Amplitudes iniciais do Setor Industrial:

0,30	0,37	0,40	0,45	0,42	0,35
------	------	------	------	------	------

Composição das anteriores com as seguintes amplitudes iniciais:

0,10 (R) + 0,40 (I)	0,15 (R) + 0,40 (I)
0,17 (R) + 0,30 (I)	0,22 (R) + 0,45 (I)
0,25 (R) + 0,42 (I)	0,20 (R) + 0,37 (I)

* Sobre estes valores de amplitude considerou-se uma variação aleatória de 0 a 10%.

Conjunto de Validação

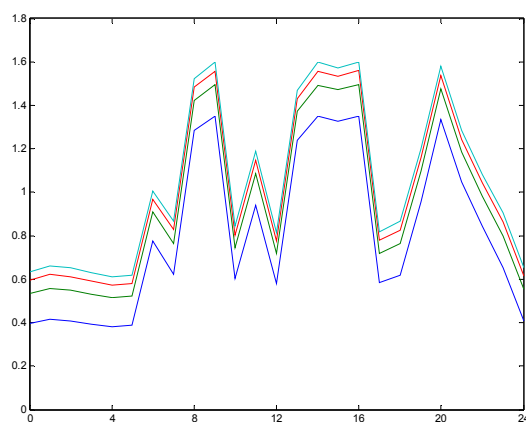


Figura 5.38: Conjunto de validação formado por composições de curvas representativas dos setores residencial e industrial submetidas a oscilações

Padrões Utilizados:

Composições de curvas referentes ao Setor Residencial e Industrial com os seguintes valores iniciais de Amplitude:

0,10 ± 10% (R) + 0,30 ± 10% (I)
0,17 ± 10% (R) + 0,37 ± 10% (I)
0,20 ± 10% (R) + 0,40 ± 10% (I)
0,22 ± 10% (R) + 0,42 ± 10% (I)

Conjunto de Validação com oscilações máximas de $\pm 10\%$ dos valores de amplitude.

Resultados Esperados		Resultados Obtidos			
		BACKPROPAGATION		AG	
<u>Saída 1</u> <i>Residencial</i>	<u>Saída 2</u> <i>Industrial</i>	<u>Saída 1</u> <i>Residencial</i>	<u>Saída 2</u> <i>Industrial</i>	<u>Saída 1</u> <i>Residencial</i>	<u>Saída 2</u> <i>Industrial</i>
0,10 \pm 10%	0,30 \pm 10%	0,1528	0,2596	0,1110	0,3384
0,17 \pm 10%	0,37 \pm 10%	0,1862	0,3574	0,1920	0,3768
0,20 \pm 10%	0,40 \pm 10%	0,2121	0,3893	0,2110	0,3864
0,22 \pm 10%	0,42 \pm 10%	0,2307	0,4097	0,2204	0,3910
SSE:		0.0053		0.0033	

Tabela 5.20: Resultados da simulação realizada com 10% de oscilações sobre as amplitudes dos padrões residencial e industrial

Conjunto de Validação com oscilações máximas de $\pm 5\%$ dos valores de amplitude.

Resultados Esperados		Resultados Obtidos			
		BACKPROPAGATION		AG	
<u>Saída 1</u> <i>Residencial</i>	<u>Saída 2</u> <i>Industrial</i>	<u>Saída 1</u> <i>Residencial</i>	<u>Saída 2</u> <i>Industrial</i>	<u>Saída 1</u> <i>Residencial</i>	<u>Saída 2</u> <i>Industrial</i>
0,10 \pm 10%	0,30 \pm 10%	0,1301	0,2985	0,0748	0,3193
0,17 \pm 10%	0,37 \pm 10%	0,2095	0,3754	0,1798	0,3555
0,20 \pm 10%	0,40 \pm 10%	0,2476	0,3991	0,2077	0,3670
0,22 \pm 10%	0,42 \pm 10%	0,2710	0,4150	0,2222	0,3729
SSE:		0.0074		0.0047	

Tabela 5.21: Resultados da simulação realizada com 5% de oscilações sobre as amplitudes dos padrões residencial e industrial

Através desta última simulação é possível reafirmar o desempenho do AG no treinamento de RNA. Pode-se perceber também que à medida que aumenta a complexidade do conjunto de treinamento proporcionalmente aumentam as dificuldades do backpropagation em encontrar o conjunto ótimo de pesos da RNA. Sendo assim se confirma a viabilidade da utilização de AG

no treinamento de RNA como alternativa ao backpropagation em aplicações de reconhecimento de composições de padrões.

CAPÍTULO VI

6 - CONCLUSÕES E TRABALHOS FUTUROS

O trabalho realizado permite conclusões relativas a cada uma das técnicas utilizadas, ao sistema híbrido proposto e à aplicação prática desenvolvida.

Com relação às RNA em específico pode-se comprovar a eficiência da técnica no reconhecimento de padrões, imunidade a ruídos e na facilidade de implementação. Uma importante questão observada ao longo do desenvolvimento das simulações é com relação às topologias das RNA. Em todos os casos foram utilizados quatorze neurônios na primeira camada e oito na camada intermediária. Qualquer alteração nestas quantidades provocou uma significativa perda na qualidade dos resultados, independente do método adotado para o treinamento e dos ajustes de parâmetros da RNA.

Outro ponto importantíssimo é definição do conjunto de treinamento. Quando se deseja associar algum tipo específico de conhecimento à RNA, amostras destes devem fazer parte do conjunto de treinamento. A quantidade de informações utilizadas para treinar a RNA não deve ser grande pois isto provocará a super especialização, e ao mesmo tempo não pode ser pequena de forma que não ocorra o aprendizado.

Com relação aos AG, foram observados pontos como o elevado custo computacional, grande número de ajustes nos parâmetros do algoritmo e as vantagens do uso desta técnica.

Como se trata de um processo aleatório, foi verificado que o sucesso do algoritmo depende de um considerável número de execuções, o que torna o processo lento.

A quantidade de ajustes de parâmetros no AG também é grande e pequenas variações em algum destes pode comprometer todo o processo.

Outra importante questão é a geração da população inicial. A forma como são gerados os indivíduos, assim como o número de indivíduos e a limitação do espaço de busca são questões essenciais ao sucesso do algoritmo.

O comprimento de cada cromossomo determina as dimensões do espaço de busca e é inversamente proporcional ao número de indivíduos. Deve-se limitar ao máximo possível o espaço de busca, isto facilita em muito a convergência do AG.

Com relação à hibridização das técnicas anteriores pôde-se comprovar a viabilidade da implementação através dos resultados apresentados. No caso específico de reconhecimento de padrões compostos observou-se que o aumento da complexidade no treinamento provocou a queda na qualidade dos resultados gerados pela RNA treinada pelo backpropagation, enquanto que o sistema híbrido manteve uma certa constância.

Isto permite verificar a consistência do sistema atuando com base na cooperação entre RNA e AG, e afirmar que nos casos onde o treinamento é complexo, é aconselhável o uso do hibridismo.

Os resultados obtidos através das inúmeras simulações realizadas, apresentados acima, são suficientes para a comprovar que o objetivo proposto neste trabalho foi atingido.

Como proposta para futuros trabalhos espera-se seja desenvolvida uma análise da defasagem entre padrões, de forma que a rede possa perceber não só alterações em amplitudes mas variações destes ao longo do eixo das abscissas.

Após esta segunda avaliação sugere-se a implementação de um sistema computacional atuando em tempo real, capaz de analisar variações nas curvas de carga e imediatamente atuar sobre a rede de distribuição de energia elétrica.

Uma ferramenta computacional como esta apenas tende agregar valor à gestão dos recursos de distribuição de energia elétrica por parte das concessionárias, além de proporcionar aumento nos lucros e conservação de energia.

CAPÍTULO VII

7 - REFERÊNCIAS BIBLIOGRÁFICAS

- [01] van Rooij, A. J. F. , Jain,L. C. and Johnson, R. P., “*Neural Network Training Using Genetic Algorithms*”, Singapore: World Scientific Publishing Co. Pte. Ltd., 1996
- [02] Vonk, E., Jain L. C. and Johnson, R. P., “*Automatic Generation of Neural Network Architecture Using Evolutionary Computation*”, Singapore: World Scientific Publishing Co. Pte. Ltd., 1997
- [03] Mesdker, Larry L., “*Hybrid Intelligent Systems*”, Kluwer Academic Publishers, 1995
- [04] Haykin, S., “*Neural Networks: a Comprehensive Foundation*”, Prentice Hall, 1999
- [05] Kovács, Z. Laszlo, “*Redes Neurais Artificiais – Fundamentos e Aplicações*”, Editora Acadêmica, São Paulo, BRASIL,1996
- [06] Minsky, M. and Papert, S., “*Perceptrons*”, The MIT Press, Cambridge, EUA, 1969
- [07] McCulloch , Warren and Pitts, Walter, “*A logical calculus of the ideas immanent in nervous activity*”, Bulletin of Mathematical Biophysics, volume 5, 1934
- [08] Hebb, D.O., “*The Organization of Behavior*”. A Neuropsychological Theory. Wiley, New York, EUA, 1949
- [09] Rosenblatt, F. , “*Principles of Neurodynamics*”, Spartan, New York, EUA, 1962
- [10] D. Rumelhart and J.L. McClelland, “*Parallel Distributed Processing: Explorations in the Microstructure of Cognition*”, Cambridge, EUA, MIT Press, 1986
- [11] Azevedo, F.M., “*Algoritmos Genéticos em Redes Neurais Artificiais*”, V Escola de Redes Neurais. pp.c091-c121. ITA, São José dos Campos, Julho de 1999.
- [12] Crestani Junior, P.R., “*Sistemas Inteligente de Navegação Autônoma: Uma Abordagem Modular e Hierárquica com Novos Mecanismos de Memória e Aprendizagem*”, Dissertação de Mestrado, Faculdade de Engenharia Elétrica e Computação da Universidade Federal de Campinas. Campinas, Dezembro de 2001
- [13] Atlas de Energia Elétrica no Brasil / Agência Nacional de Energia Elétrica – Brasília: ANEEL, 2002
- [14] Goldberg, David E., “*Genetic Algorithms in Search, Optimization and Machine Learning*”, Addison-Wesley Publishing Company, Inc., 1989.

[15] Holland J.H., "*Adaptation in Natural and Artificial Systems*", MIT Press, Cambridge, Massachusetts, 1975.

[16] Montana, D.J., "*Neural Network Weight Selection Using Genetic Algorithms*", Bolt Beranek and Newman Inc., 70 Fawcett Street, Cambridge, MA 02138.

[17] Tanomaru, J., "*Motivação, Fundamentos e Aplicações de Algoritmos Genéticos*", II Congresso Brasileiro de Redes Neurais, III Escola de Redes Neurais, Curitiba, 1995

[18] Nunes, L.E.N.P., Prado, P.P.L., "*Reconhecimento de Objetos Contidos em Imagens através de Redes Neurais*", Rev. Ciênc. Exatas, Taubaté, v.5-8, p.77-88, 1999-2002

[19] Todesco, J.L., Pimentel F.J.S., Bettiol, A.L., "*O Uso de Famílias de Circuitos e Rede Neural Artificial para Previsão de Demanda de Energia Elétrica*", Anais do XXIV Encontro Nacional de Engenharia de Produção, Florianópolis, SC, BRASIL, 2004

[20] Lazo, Juan G.L., "*Sistema Híbrido Genético-Neural para Montagem e Gerenciamento de Carteiras de Ações*", Dissertação de Mestrado, Departamento de Engenharia Elétrica PUC/RIO, Rio de Janeiro, 2000

[21] Matsumoto, E.Y., "*MATLAB 6.5: Fundamentos de Programação*", Editora Érica, São Paulo, 2002